A Pathway to Mobility: Shifting Low-Income Students to Higher-Quality Degrees

Elizabeth Burland, Nora Delaney, Susan Dynarski, Katherine Leu, CJ Libassi, Katherine Michelmore, Stephanie Owen, Mary Quiroga, and Elizabeth Salinas*

Abstract

Earnings inequality is primarily growing within education groups, especially among college graduates. Observed returns to a bachelor's (BA) degree vary substantially by college. Disadvantaged students are concentrated at schools with lower payoffs, undermining the potential of college to act as an engine of income mobility. We experimentally test a program to shift low-income students toward higher quality degrees. Students randomly offered a guarantee of four years of free tuition at the flagship University of Michigan-Ann Arbor were about twice as likely to earn a BA from a highly selective college. LATE estimates indicate that those induced to attend the flagship are 65 percentage points (pp) more likely to earn a BA of such high quality, relative to a control complier mean of 4 percent. The offer increased four-year BA completion by 4 pp, with a 23 pp increase implied for compliers. After six years, the treatment and control groups are equally likely to have a BA. We find no decrease in the likelihood of graduating with a STEM degree. The shifts in degree quality imply large increases in income mobility. LATE estimates predict a 31 percent increase in earnings at age 30 for those induced to attend a higher quality college.

---

*Burland: University of Connecticut; Delaney, Dynarski (dynarski@g.harvard.edu), and Salinas: Harvard University; Leu, Michelmore, and Quiroga: University of Michigan; Libassi: Teachers College, Columbia University; Owen: Colby College. We thank collaborators at the University of Michigan Office of Enrollment Management, particularly Kedra Ishop, Steve Lonn, Paul Robinson, and Betsy Brown. We are grateful for research assistance from Zia Kinzy and feedback from Chris Walters and many seminar participants. This study is registered as RCT ID AEARCTR-0001831, with 10.1257/rct.1831-4.0. A pre-analysis plan was initially filed in April 2017. This research used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data is modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). The research was supported by IES grant R305B150012 and the Smith Richardson Foundation. This research does not represent the views of IES or the US Department of Education.

I. Introduction

Over the late 20th century, increasing returns to education drove rising earnings inequality. In the 21st century, by contrast, earnings inequality has primarily grown *within* education groups, especially those with a bachelor's (BA) degree (Autor, Goldin, and Katz, 2020). Growing evidence suggests that the quality of a college drives the economic returns its degree delivers. Low-income students disproportionately attend open-access and less selective schools, while high-income students are far more likely to enroll in highly selective colleges. In the Ivy League, there are more students from the top one percent of the parental income distribution than there are students from the bottom 50 percent (Chetty et al., 2020). These enrollment patterns raise concerns that postsecondary education in the US may act more as a reinforcer of class rigidities than as an engine of economic mobility.

In Michigan, differences in income mobility between schools are stark. Half of low-income students who attend the University of Michigan flagship in Ann Arbor (hereafter, UM flagship) reach the top income quintile as adults, compared with just 10 percent at Grand Rapids Community College, 19 percent at Eastern Michigan University, and 34 percent at Michigan State University[1] ("low-income" in the cited statistics is defined as the bottom quintile of parental income). But while the UM flagship does very well at increasing upward mobility among low-income students, it enrolls few of them: as of 2014, just three percent of its undergraduates were from the bottom quintile of parental income. Three times as many (nine percent) were from the top *one percent* of the income distribution (Chetty et al., 2022).

These mobility patterns do not identify the causal effect of attending a more selective college. Students select institutions based on family resources and academic preparation, so simple comparisons confound selection with treatment. Quasi-experimental designs using admissions cutoffs and waitlists identify sizable returns to college quality for students at the margin of admission. Less is known about the broader set of low-income students who could plausibly attend highly selective colleges. We provide the first experimental estimates of the effect of attending a highly selective college.

We fielded a randomized intervention designed to encourage enrollment of low-income students at Michigan's highly selective flagship. The High Achieving Involved Leader (HAIL) Scholarship guaranteed to high school students four years of free tuition and fees at the flagship if they were admitted. Low-income students were pre-qualified using data on subsidized lunch eligibility from state administrative records; no financial aid forms, including the FAFSA, were required. This was not an increase in aid, since the flagship already fully met financial need.

[1] The colleges we use throughout the paper as examples are frequently attended by students in the control group.

Rather, it was a retiming of when the aid offer was made, from after admission to before applications.

HAIL more than doubled the rate of application and enrollment at the UM flagship, diverting students from community colleges and less selective four-year schools (Dynarski et al., 2021). Because the scholarship did not change aid received conditional on enrollment at UM, random assignment of the offer generates clean variation in enrollment at a highly selective flagship for high-achieving, low-income students.

We estimate intent-to-treat effects of the offer on degree attainment, degree quality, and field of study. HAIL accelerated completion, raising four-year BA attainment by 4 percentage points (11 percent), with no effect on BA attainment after six years. It also upgraded degree quality: treated students were 12 percentage points (78 percent) more likely to earn a BA from a highly selective institution. The treatment did not affect the share of students earning STEM degrees; treated students are just as likely as the control group to earn a STEM degree.

We use the exogenous variation in college quality created by the experiment to estimate local average treatment effects. The offer increased enrollment at the UM flagship within two years of high school by 18 percentage points. Attending the flagship speeds completion: treated compliers are 23 percentage points more likely to earn a BA within four years than control compliers, whose on-time completion rate is 41 percent. The intervention induced a substantial increase in the quality of BAs earned. Enrolling at the flagship increases by 65 percentage points the likelihood of earning a BA from a college as selective as the UM flagship.[2] This is extremely large relative to the control-complier mean of four percent.

We show that HAIL diverted students toward the UM flagship from a wide range of counterfactual colleges, from open access community colleges to less-selective four year schools. Attending the UM flagship represents a much larger increase in college selectivity for a student who would have otherwise attended a community college than for a student who would otherwise have attended a four-year college. To provide context for our estimates, we show how the distribution of students' college quality was altered by the intervention.

Based on this increase in BA quality, we project substantial increases in earnings. As the sample is too young to obtain meaningful estimates at the individual level, we project earnings using college-level data from the Internal Revenue Service. Attending the UM flagship is predicted to boost earnings at age 30 by 31 percent (an increase of $17,800 relative to a control complier

[2] Barron's classifies the UM flagship as a "highly competitive" university, the second highest category.

mean of $57,000). This is equivalent to moving a graduate from the forty-sixth to the sixty-third percentile of the US earnings distribution for those with a BA.

II. Motivation and Prior Literature

II.A. What is college quality?

"College quality" has no formal definition. "High quality," "highly selective," and "elite" are used interchangeably by researchers, the media, and policymakers. A group of correlated characteristics including selectivity, peer achievement, student-faculty ratios, and institutional resources predict a college will have higher graduation rates and higher graduate earnings (Black and Smith, 2006), which are the outputs generally used to measure quality.

Higher quality schools have faster rates of degree completion. In the US, an on-time bachelor's degree typically requires four years of study, with thirty college credits earned each year. But many students take considerably longer, especially at less-selective colleges.

Nationwide, only 48 percent of students graduate within four years (U.S. Department of Education, National Center for Education Statistics, 2024). Reflecting that graduating in more than four years is the norm, the federal website devoted to tracking college performance uses six-year graduation as its primary metric (College Scorecard, 2023).

In Michigan, the four-year completion rate is 81 percent at the UM flagship, 59 percent at Michigan State University and 41 percent at Grand Valley State University (U.S. Department of Education, National Center for Education Statistics, n.d.).[3] Using the California system as a point of reference, these differences are similar in magnitude to those between the University of California-Berkeley, University of California-Davis, and California State Long Beach.

A lengthy time-to-degree is costly to both students and society. Tuition is charged by the year, not by the degree. The opportunity cost, in the form of forgone wages, accrues annually. Finally, the later in life a degree is granted, the fewer years there are for it to deliver returns. For all these reasons, researchers and policy actors have focused intently on the determinants of time-to-degree and interventions to reduce it (e.g., Bound et al., 2012).

In addition to having better outcomes and more resources, higher-quality colleges tend to have more generous financial aid. The University of Michigan flagship is the cheapest four-year

[3] Differences in graduation rates shrink after six years of enrollment, to 94 percent at the UM flagship, 83 percent at Michigan State University, and 67 percent at Grand Valley State University (College Scorecard, 2023).

option in the state for low-income students (Dynarski et al., 2021). The only similarly priced options are local community colleges.

We turn next to evidence on the causal effects of college quality on educational attainment and earnings.

II.B. Prior evidence on the causal effect of college quality

It is difficult to estimate the effect of college quality on student outcomes. Students sort into colleges based on characteristics that likely have their own effects on college success. Untangling the effect of those student characteristics from that of college quality is an empirical challenge. Further, counterfactuals vary widely across studies, which makes it difficult to compare results. We summarize the existing research in Appendix Table 1.

In one of the early quasi-experimental estimates in this literature, Dale and Krueger (2002) compare outcomes for applicants applying to and accepted by similarly selective schools (e.g., Princeton and Yale). They conclude (using self-reported earnings data from students who started college in 1976) that college selectivity has little effect on earnings. An update to that paper (using administrative earnings data and adding a cohort that started college in 1989) finds that attending a more selective school increases earnings but only among Black, Hispanic, and first-generation students (Dale and Krueger, 2014).

More recent research exploits discontinuities in admissions cutoffs to estimate the causal effect of college quality and generally finds positive effects on college graduation and earnings (see Lovenheim & Smith, 2023 for a review). Zimmerman (2014), Goodman, Hurwitz, and Smith (2017), and Kozakowski (2023) estimate the effect of gaining admission to a four-year public school in contexts where the counterfactual is a community college or no college. They find positive effects on six- and eight-year bachelor's degree completion. Mountjoy (2026) exploits cutoffs for public, four-year schools in Texas, where the counterfactual is a mix of community colleges and other four-year schools. He finds positive effects on bachelor's degree completion and earnings.

Chetty et al. (2026) use a quasi-experimental design to estimate the effect on earnings of attending the most selective colleges in the US, the "Ivy-plus" schools. Students admitted to these elite colleges are more likely to reach the top one percent of the income distribution and work in elite occupations than students who attended a state flagship. It is worth noting here that "state flagship" spans a wide range of selectivity. The UM flagship sits at the upper end of that category, with peer achievement and resources closer to Ivy-plus colleges than to a typical state flagship.

Experimental evidence also points to positive effects of college quality, but relies on interventions that both affect college choice and provide extra academic or financial support to

students. Angrist, Autor, and Pallais (2022) evaluate the impact of a scholarship program that provides generous financial aid to graduating high school students from low-income families. Scholarships were randomly assigned among a subset of students who applied for a scholarship with modest academic requirements. The scholarship shifts students away from community colleges into four-year colleges. Among students who had indicated a preference for four-year colleges, six-year BA attainment rates rose eight percentage points. Scholarship beneficiaries continued to receive grant aid from the sponsoring foundation, beyond what was provided to their classmates. The evaluation of this intervention therefore captures both the effect of college quality and the effect of additional financial support during college.

Barr and Castleman (2025) evaluate the impact of pre- and in-college advising on college choice and degree attainment. Students randomly assigned to college advising were eight percentage points (16 percent) more likely to obtain a bachelor's degree after six years. While the authors conduct several tests that suggest that the effects are largely driven by shifting students into higher quality colleges, they cannot rule out direct effects of the college advising itself, which continues throughout college.

Cohodes, Ho, and Robles (2024) use a mix of experimental and quasi-experimental methods to examine a STEM-focused summer program hosted by an elite university, which induced very high-achieving, minority students to enroll in very selective schools. The summer program provided mentorship and academic support to targeted high school students. Their results are imprecise but suggest that the program increased on-time graduation by three to seven percentage points and BA attainment by two to nine percentage points. Treated students were 11.5 percentage points more likely to earn a degree from an elite school.

For completeness, we note a small literature that finds negative effects of attending a more selective college. Light and Strayer (2000) conclude that students in the lowest quartile of academic preparation are less likely to graduate if they enroll at higher-quality colleges. But Dillon and Smith (2020), using the same data, show that students of all preparation levels benefit from higher quality schools.

Arcidiacono et al. (2014) conclude that higher-ranked schools do worse at graduating less-prepared students. They find that after California's affirmative action ban shifted minority students into less selective schools, their graduation rates increased, they were more likely to earn STEM degrees, and they graduated more quickly. However, Bleemer (2022) finds the opposite: graduation rates, STEM degrees, and earnings are higher for students admitted to higher quality schools through affirmative action.

A qualitative literature studies the experiences of low-income students at elite colleges. It finds that low-income students have difficulty navigating the norms of selective colleges (attending office hours, obtaining extensions, requesting letters of recommendation) and meeting basic

needs such as food, housing, and transportation (Goldrick-Rab, 2016; Jack, 2019; Jack & Black, 2024). While this literature provides evidence for why low-income students might face more challenges navigating elite colleges than their higher-income classmates, it provides no evidence of how college quality affects the educational outcomes of low-income students.

II.C. Contribution

We address two important gaps in this literature.

First, the literature focuses almost exclusively on academically marginal students. The quasi-experimental evidence that relies on admissions cutoffs estimates the effect of quality for students whose academic achievement is (by definition) far below that of their college classmates. This literature therefore confounds the effect of college quality with the effect of low class rank. By contrast, the students in our sample have SAT scores typical at the UM flagship, scoring higher than 85 percent of SAT-takers nationwide. We therefore provide experimental evidence on the effect of college quality for students who are academically typical at their school.

Second, the thin experimental evidence on college quality is from interventions that provide to treated students financial and/or academic support that plausibly has its own effect on student success. These studies therefore confound the effect of increased college quality with the effect of additional financial and academic support. As we next describe, the HAIL intervention induced treated students into a high-quality college but it provided no additional support to enrollees beyond that typically provided by the college.

III. The HAIL Scholarship

III.A The Intervention

We partnered with the UM flagship to increase its enrollment of low-income students. We jointly designed and implemented a randomized intervention,[4] the HAIL Scholarship. HAIL offered an up-front guarantee of four years of tuition and fees at the UM flagship to selected students. The offer conferred no admissions advantage. Students in both treatment and control applied through the standard admissions process, which reviews transcripts, test scores, essays, letters of recommendation, and extracurricular activities.

[4] The randomized trial was approved by the IRB at the University of Michigan (HUM00096289). The present analysis is overseen by the University of Michigan (HUM00224676), Harvard University (IRB21-1221), Colby College (2021-093), and the University of Connecticut (X23-0649).

We identified eligible students using state administrative records. Students qualified if they were eligible for subsidized school meals and performed above test-score and GPA cutoffs chosen by the UM admissions office.[5] Because Michigan has free, in-school, mandatory SAT testing for 11th graders in public schools, the intervention was able to target low-income students who, in other states, may not have attempted the SAT.[6] Qualifying SAT scores start at 1100 while qualifying GPAs start at 3.3. Students with higher test scores faced a lower GPA threshold (and vice versa).[7]

Applicants from treatment and control schools were admitted to the flagship at similar rates. Admission rates for HAIL-eligible applicants were higher than in the broader applicant pool, where the admission rate is 29 percent (Dynarski et al., 2021). This suggests that our intervention did not induce marginally qualified students to apply; if anything, they were more qualified than the typical applicant.

The intervention consisted of targeted outreach that communicated the guarantee. Students in the treatment group were mailed a packet describing the HAIL offer and encouraging them to apply. Parents received a follow-up letter with the same information, while high school principals received a list of eligible students and information about the scholarship (see Dynarski et al., 2021 for materials and implementation details). Control schools received the UM flagship's standard outreach materials, which emphasized the school's generous scholarships and high academic standards but made no concrete promise of financial aid.

HAIL differed in several key dimensions from typical financial aid. In the traditional system, students learn net price only after applying and being admitted to a given college. HAIL offered an aid guarantee before college applications were due. In the traditional system, students reapply annually for aid, which can change from year to year. HAIL provided a four-year commitment in the fall of senior year and required no financial aid application forms. By replacing uncertainty and administrative burden with an early, credible guarantee, the offer changed enrollment choices among high-achieving students from low-income families (Dynarski et al., 2021; Burland et al., 2023).

HAIL did not increase aid, but rather shifted when it was promised. Among those who attended the UM flagship, students in the treatment and control groups received the same amount of grant aid both annually and cumulatively (Appendix Table 2). The two groups also received the same

[5] The family income threshold for subsidized meal eligibility is 185 percent of the federal poverty line.
[6] Hyman (2017) shows universal testing induced disadvantaged students to take the college admissions test, with many newly appearing in the upper end of the score distribution.
[7] Michigan switched from the ACT to the SAT in 2016, one year into the study; we convert ACT scores to SAT using a crosswalk.

academic and social services at the flagship, with no special resources devoted to HAIL admits. This supports our exclusion restriction in the LATE analysis: the HAIL offer affects educational attainment and adult earnings only through the channel of improved college quality.

We learned of an additional implementation detail after the publication of our first two studies of the intervention's impacts (Dynarski et al., 2021, Burland et al., 2023): the admissions office made an offer of deferred admission to some of the treatment group. Affected students were promised transfer to the UM flagship on the condition they completed a year at a UM satellite (Dearborn or Flint), with the tuition guarantee applying to this pathway.[8] The satellite campuses have considerably lower peer achievement and graduation rates than the flagship,[9] so the deferred admission offer could, in theory, reduce college quality for students diverted from other four-year colleges to a UM satellite. We explore this in Section VI, where we show that the satellites, while of lower quality than the flagship, are of similar quality as the colleges attended by the control group. A comparison of the distributions of college characteristics for treatment and control suggests that the intervention generated a monotonic increase in college quality.

III.B Randomization

Random assignment occurred at the high school level, with all eligible students at a given school assigned the same treatment status. School randomization occurred within four strata defined by the number of HAIL-eligible students in a school. High schools with eligible students in both years of the experiment maintained their treatment status in the second year.

We study 3,910 students who graduated from high school in 2016 and 2017: 1,932 treated students (from 289 unique schools) and 1,978 control students (from 299 schools). Table 1 shows the experimental balance on school-level characteristics. There are no significant differences in these characteristics by treatment status. In the control group, the average SAT score is 1254 and the average GPA is 3.82. Fifty-three percent of control-group schools are in rural areas, 40 percent are in Southeast Michigan, and average 93.2 miles from the University of Michigan. All students are eligible for free (71 percent) or reduced-price (29 percent) lunch. The sample is primarily white and non-Hispanic and majority female.

[8] In the first cohort, 147 treatment students were offered deferred admission to the flagship if they enrolled at Dearborn. Eighty students (54 percent) took this offer. In the second cohort, 214 treatment students were offered deferred admission to the flagship if they enrolled at Dearborn or Flint. Seventy-five students (35 percent) took the offer. Over the two cohorts, sixty percent of students who took the offer enrolled at Dearborn and 40 percent at Flint.
[9]The average SAT score for in-state students is 1360 at the flagship, 1160 at Dearborn, and 1062 at Flint. Six-year graduation rates are 81 percent at the flagship, 56 percent at Dearborn, and 36 percent at Flint (U.S. Department of Education, National Center for Education Statistics, 2024).

IV. Data and Methods

IV.A. Data

We obtained longitudinal, student-level administrative data from the Michigan Center for Educational Performance and Information (CEPI) that contain the universe of students attending Michigan public schools. The distribution of student achievement in Michigan mirrors the national distribution, with students in 2017 averaging around 1000 on the SAT with a standard deviation of 195 (U.S. Department of Education, National Center for Education Statistics, 2017b). Our sample is limited to Michigan students who are both low-income (certified eligible for federally subsidized school meals) and high-achieving (meeting high school GPA and test score cutoffs set by UM's admissions office).

Our primary educational outcome is degree attainment, measured within four and six years after expected high school graduation. We categorize degrees as "highly selective" if they are considered highly competitive or most competitive by the Barron's selectivity index. We also examine the field of degree and credits earned. Detailed information on how we define our outcome measures is included in Appendix B.

To track postsecondary enrollment and degree completion, we link to National Student Clearinghouse (NSC) data, which provide college attendance and credential records from virtually all U.S. colleges and universities. NSC data includes the field of degree. We supplement NSC records with transcript data from the Student Transcript and Academic Record Repository (STARR), which include information on credits attempted and earned at each college.[10] Data on college characteristics are drawn from the National Center for Education Statistics and College Board.

The STARR transcripts include only students in public colleges in Michigan. Since the intervention shifts enrollment toward Michigan's public colleges, we need to attend to selection into STARR. We take two approaches, detailed in Appendix B. First, in our main analyses we impute missing credits using the approach of Angrist, Autor, and Pallais (2022), predicting credits based on selectivity and sector of college attended. Second, as a robustness check, we calculate Lee (2009) bounds. The results from the two approaches are consistent in their conclusions on effect sizes, though the Lee bounds are wide.

Data on adult earnings are not yet available for our cohorts. Publicly-available tax data provide median earnings at age 30, by college, for cohorts close in age to our sample (Chetty et al.,

[10] We do not reliably observe credits transferred into a college, whether from advanced high school classes or from other colleges.

2022). We impute earnings based on the modal college that students attend during the four years after high school. All dollar amounts are converted into 2025 values.

IV.B. Empirical Strategy

We estimate the effects of the HAIL Scholarship on degree attainment, degree selectivity, STEM degree completion, credit accumulation, transfers, and predicted earnings. Because treatment was randomly assigned at the high school for each cohort, our analytic unit is the high school-by-cohort cell:

$$(1)\ Y_{jt} = \beta_0 + \beta_1 HAIL_j + S_j + u_{jt}$$

Where Y_{jt} is an outcome of interest for cohort t at high school j. $HAIL_j$ indicates assignment to the treatment group, and S_j is a vector of strata indicators. We cluster standard errors at the high school level. β_1 measures the causal effect of being offered the HAIL Scholarship (ITT).

To estimate the causal effect of college quality, we compute LATEs using instrumental variables (IV). We use the randomized HAIL offer as an instrument for enrolling at the UM flagship within two years of high school graduation. The LATE captures the effect of upgrading college quality for students whose enrollment decisions were altered by HAIL. We estimate the following with two-stage least squares, with UM_{jt} representing the rate of enrollment at the flagship for cohort t from high school j and γ_1 representing the LATE.

$$(2)\ UM_{jt} = \alpha_0 + \alpha_1 HAIL_j + S_j + e_{jt}$$
$$(3)\ Y_{jt} = \gamma_0 + \gamma_1 \widehat{UM_{jt}} + S_j + u_{jt}$$

Our exclusion restriction requires that HAIL affects downstream outcomes only through changes in the college attended. We discuss the plausibility of this restriction and the potential role of always-takers in Section VII.C.

V. Intent-to-Treat Effects

V.A. Initial Enrollment

The estimated effect of HAIL on enrollment forms the first stage for the IV estimates and defines the complier counterfactual. The HAIL offer increased the likelihood of enrolling at the UM flagship within one year of high school by 15 pp (Table 2, Figure 1a). The treatment-control contrast suggests that about a third of students induced into the flagship would have otherwise attended a two-year college (4 pp). Two-thirds would have attended a four-year college that is

less selective than the flagship (11 pp). The offer had no effect on college entry or on enrollment at colleges at least as competitive as the UM flagship.[11]

The deferred-admission offer induced six percent of students to enter UM-Dearborn or UM-Flint as freshmen (Figure 1b), a large effect. Half that number subsequently transferred to the flagship the following year (+2.9 pp, Table 3). Transfers from the satellites to the flagship are absent in the control group (0.001 in year two, Table 3), suggesting that the offer of deferred admission indeed expanded access to the flagship. This set of estimates supports the monotonicity assumption that underpins the LATE analysis.

The first stage is larger when we expand the time frame to include the second year, reflecting the deferred admissions offer. The offer increased flagship enrollment by an additional three percentage points in year two, for a total increase of 17.8 percentage points within two years of high school (standard error 1.9 pp, F-statistic 85). We use this more expansive definition of enrollment in the flagship in defining the first stage for our IV estimates.

V.B. Degree Completion and BA Selectivity

We now examine how the HAIL offer affected bachelor's degree attainment and the selectivity of the degree-granting school. The offer increased four-year BA completion by 4.1 pp (Table 2, standard error 2.2 pp), relative to a control mean of 39 percent. While the estimate is only marginally significant, we show below that the transcripts (which come from a different data source) corroborate this finding.[12] By six years, BA attainment is 62 percent in both groups.

While the likelihood of BA receipt is unaffected by the HAIL offer, we find large, persistent and positive effects on degree quality. Within four years, treated students are 11.3 pp more likely to earn a BA from a highly- or most-selective university. The control mean is 10.3 percent, indicating the HAIL offer doubled the likelihood of graduating with a BA this selective. By six years, the effect has risen to 11.7 pp (control mean is 14.4 percent). All of these estimates are statistically significant at conventional levels. Within the treatment group, 43 percent of BAs are from a highly selective college, compared to 24 percent of BAs in the control group. As we show later, this predicts substantial earnings gains later in life.

We find no effect on the likelihood that a student's BA is in a STEM field (Table 2). The undermatch hypothesis posits that disadvantaged students induced into more competitive academic environments will avoid rigorous majors (Arcidiacono et al., 2014). We find no evidence supporting this hypothesis. A quarter of both treatment and control students earn a

[11] Greater detail on the enrollment effects of HAIL is available in Dynarski et al. (2021) and Burland et al. (2023).
[12] The enrollment and degree results in the paper are estimated using data from the National Student Clearinghouse, while credit data is drawn from STARR.

STEM BA. This finding is unchanged with alternative STEM definitions, such as adding business and economics to the STEM group (Appendix Figure 2).

In sum, the HAIL offer changed degree quality and timing but not overall BA attainment or the field of degree. We next examine mechanisms that may underpin shortened time to degree: faster credit accumulation and decreased enrollment at community colleges.

V.C. Faster Credit Accumulation

Nearly all students at the UM flagship enroll full time. Three percent attend part time, compared with nine percent at Michigan State University and 68 percent at Grand Rapids Community College (College Scorecard, 2016). These differences in enrollment intensity suggest a potential mechanism for shortened time-to-degree, since a full-time student will tend to accumulate credits faster than a part-time student.

Indeed, students in the treatment group accumulate credits at four-year colleges more quickly than those in the control group (Figure 2, Appendix Table 3). This starts in the summer after high school, with the treatment group earning 0.34 (standard error 0.09) more four-year credits than the control group. The UM flagship has a pre-college summer program for students from disadvantaged backgrounds, which likely explains this result. The credit advantage grows with time, to 1.3 after the freshman year, 1.98 after the sophomore year, and leveling off at about 2.5 after that. While the estimates for the first two spells are statistically significant at conventional levels, the confidence bands include zero for the later years.

These differences in credit accumulation may seem small, but they correspond to substantial differences in the share of students who are on-time to earn a degree within four years. The typical BA in the United States requires 120 credits. Students therefore need to enroll in 30 credits per year in order to graduate on time.

By the end of the first year of college, treated students are five percentage points more likely to have earned 30 credits at a four-year school (relative to a control mean of 31 percent, Appendix Table 3, Panel B). Treated students are three to five percentage points more likely to reach the 30-credit benchmark over four years of college. This is in line with the estimated effect on four-year BA completion (4.1 pp, Table 2), pointing to faster credit accumulation as a central mechanism for the shift toward on-time graduation. This pattern aligns with evidence that early credit accumulation strongly predicts degree completion (Angrist et al., 2022).

In Appendix Table 3, we report credits earned at two-year colleges. Treated students earned fewer credits in the two-year sector, reflecting reduced community college enrollment. The decline in community college credits is smaller than the gain in four-year credits, so total credits earned are increased by the treatment.

Credit information is from the STARR data, which covers only students in Michigan's public colleges. In the first year after high school, there is no difference in the NSC data between the treatment and control group in the share enrolled in any college, 90 percent (see Figure 1A). But in the same time frame we observe 73 percent of the control group in the STARR data, compared to 77 percent of the treatment group (see Appendix Table 5). The difference, four percentage points, is the rate of selection into STARR induced by the HAIL offer (the first year is used as an example; the rate of selection for each year is shown in Appendix Figure 3). These students would have otherwise attended a private college or a college outside of Michigan, but were induced by the HAIL offer to attend a public college in Michigan.

This creates a credible threat of selection bias in our credit estimates, since presence in the STARR data is selected on an outcome downstream of random assignment. As described in the data section, we imputed credits for such students for Figure 2 and Appendix Table 3. As a robustness check, we estimated effects with the unimputed sample and calculated Lee bounds on the resulting parameters (see Appendix section B.6). The Lee bounds are wide (see Appendix Figure 4) but accord with the analysis based on imputed credits.

V.D. Reduced Attendance at Two-Year Colleges and AA Degree Attainment

Community college tuition is well below that at four-year colleges (College Scorecard, 2016). Online advice on how to minimize college costs often recommends that students first enroll in a community college and then transfer to a four-year college to earn a BA.[13] The largest state aid program for low-income students in Michigan *requires* that students first enroll in a community college and earn an AA before they can use the scholarship at a four-year college (Burland et al., 2025).

Evidence suggests that this choice can actually harm students, through slower progress and credit loss during transfer (Miller, 2025). In the United States, just 18 percent of those starting at a community college have earned an AA after six years (U.S. Department of Education, National Center for Education Statistics, 2017a). Even fewer (13 percent) earn a BA (U.S. Department of

[13] For example, the state of Michigan's "Planning for your child's college education" (https://www.michigan.gov/financialfuture/toolkit/planning-for-your-childs-college-education) as well as nation-wide resources including Liberty Bank (https://www.liberty-bank.com/about-liberty-bank/learning-center/2026/01/05/8-ways-to-make-college-more-affordable) and the College Raptor (https://www.collegeraptor.com/paying-for-college/articles/financial-advice-planning/college-affordability-how-a-community-college-can-fill-the-gap/), mention starting at community college as one strategy to keep college costs down. Further, the AI-assisted results from a Google search querying "how to minimize college costs" returns a list of bullet-pointed strategies, the first of which is "Community College & Transfer: Complete general education requirements at a lower-cost community college before transferring to a four-year institution." (accessed February 23, 2026).

Education, National Center for Education Statistics, 2017a). We observe this slow progress in the control group: while 13 percent start out at a community college (Figure 1), after one year only 2.9 percent have earned enough credits to be on-time for an AA (Appendix Table 3).[14]

As shown earlier, the HAIL offer reduced the share of students starting out at a community college by four percentage points (Figure 1, Panel A). The HAIL offer decreased the share of students earning an AA by six percentage points (relative to a control mean of 15 percent, Figure 3, Panel B). The decline reflects both reductions in the probability of earning only an AA (three pp) and of earning both an AA and a BA (three pp). There is an increase in the share earning only a BA (two pp; Figure 3, Panel B) and a decline in the share of students earning any degree (three pp).

The decrease in AAs combined with no change in BAs is consistent with two scenarios, with differing welfare implications. Some students who were diverted from community college to the UM flagship may otherwise have earned an AA and then gone on to earn a (less-selective) BA. For these students, earlier enrollment at the flagship both saves time and increases BA selectivity, predicting higher lifetime earnings.

But some students diverted from community colleges to the UM flagship may have otherwise earned an AA, and ended up instead with some four-year credits and no BA. The welfare implications of this second scenario depend on the return to an AA as compared to the return to four-year credits with no degree (note that total credits earned in the treatment and control group are the same). In later research, we plan to examine employment outcomes to try to pin down these welfare implications.

V.E. Transfers

The pattern of transfers we observe in the data is consistent with the findings discussed above. HAIL decreased transfers from community colleges into four-year colleges in the third year after high school (by 1.4 pp, Table 3). This pattern suggests that some students induced into UM would have otherwise spent two years in a community college before transferring to a four-year school.

The HAIL offer also increased transfers from the UM satellites to the flagship by 2.9 percentage points in the second year, reflecting take-up of the deferred-admission offer. Transfers from the

[14] Note that this statistic reflects credits *earned* at a two-year college. More than this may have *accumulated* at least 30 credits toward a two-year degree, through transfer of college-level coursework from high school. We do not reliably observe transfer of credits between schools.

satellite campuses to the flagship are essentially absent in the control group, suggesting that the deferred-admission offer opened a new path into the flagship.

We see a small increase in transfers away from the flagship. In the second year after high school, transfers from the flagship to colleges outside the UM system rose by 1 percentage point. Recall that HAIL increased first-year enrollment at the flagship by 15 percentage points. Therefore, one out of every 15 students induced to enroll in the UM flagship in their first year left for another college during their second year.

VI. Induced Changes to College Quality

The existing estimates of the effect of quality on downstream outcomes, while generally positive, vary widely in magnitude. For instance, attainment effects range from null to 8 percentage points; earnings effects range from 0 to 14% (see Appendix Table 1). We could get such a range of estimates if the *effect* of quality is stable but its *dosage* varies considerably across the contexts under study. A shift from a community college to a four-year college will plausibly generate different effects than a shift between four-year colleges of differing quality.

In our context, we have unusually detailed information about counterfactual colleges and the students that attend them. We use these data to explore both the mean increase of college quality induced by HAIL and how it varies across counterfactuals. To capture the effect of transfers, including those induced by the offer of deferred admission, we calculate the average quality of the colleges a student attends in the first two years after high school.

VI.A. Peer Achievement and Academic Match

We start by examining the academic preparation of college peers. We quantify peer quality by calculating the average SAT score of a student's college classmates. We can make this calculation for all colleges, not just the selective schools that require admissions testing, because the SAT is universally required in Michigan's public high schools. We therefore can characterize peer quality for open-access institutions, which are typically impossible to place on a common score scale with the more selective schools that require the SAT. We standardize to a mean of zero and a standard deviation of one using values for the college peers of the control group.[15] Students who do not enroll in college are in their own category and have the lowest average scores.

In Figure 4 we show the distribution of peer quality for the treatment and control groups, normalized by the control group. The UM schools are marked, as are a few other schools

[15] We use SAT scores for Michigan's public-school students to calculate the college-level averages.

frequently attended by the control group. The UM flagship is a significant outlier in peer achievement, with peer scores 1.7 standard deviations above the colleges typically attended by the control group. A very small handful of students attend schools with peer quality higher than that of the UM flagship.

The UM satellites bracket the mean for the control group, with Flint slightly below the mean (-0.45sd) and Dearborn slightly above (0.23sd). This is evidence that these satellites were typical of colleges attended by the control group, rather than unusually weak alternatives. The rightward shift is consistent with a monotonic effect of the HAIL offer on peer quality, with no increase in mass at or below the satellite campuses.

As is clear from the figure, the intervention strongly shifted the distribution of peer quality to the right, creating a large mass at the flagship. The treatment group is missing mass among a wide range of counterfactual schools, indicating heterogeneous effects of the treatment on peer achievement.

As peer achievement shifts so too does the academic match between a student and their peers. Figure 5 plots the distribution of each student's SAT score relative to the average SAT score at their college. Positive values indicate that a student's SAT score exceeds the peer mean. Typically, low-income students lag behind their college peers in their achievement scores. In Michigan, 66 percent of low income students attend colleges where their SAT score is below the peer mean (Figure 5, Panel A).

The HAIL sample is not typical, with a mean SAT score at the 85th percentile of the national distribution. Control students attend colleges where their score is well above the peer mean (Figure 5, Panel B). Eighty-two percent of control students have SAT scores above their college peers' average. On this metric, students in our sample are undermatched, not overmatched. The HAIL offer reduces undermatch: in the treatment group, 75 percent of students score above the mean of their peers. The HAIL sample is very different from the compliers that identify effects in the quasi-experimental literature on college quality: marginal admits scoring well below their college peers.

VI.B. Institutional Quality

We now turn to how HAIL affected other dimensions of college quality, such as instructional spending, endowments, admissions rates, and student-to-faculty ratios. We also examine peer demographics, including parental education, Pell Grant receipt, and family income. We code these variables so that higher values indicate higher college quality.

Figure 6 plots the distributions for peer characteristics and institutional characteristics by treatment status; we report average treatment effects in Appendix Table 4. Across measures, the

offer shifts students toward colleges with characteristics typical of more selective colleges. The treatment-control contrast is clearly visible for each metric.

Students offered HAIL shift away from schools with admission rates of 50 to 100 percent toward the flagship, which admits about 29 percent of applicants (Figure 6a).[16] Control students typically attend schools with instructional expenditures of $8,000 to $21,000 per student, compared to over $30,000 at the flagship (Figure 6b). The endowment distribution is particularly skewed: the flagship has an endowment of over $300,000 per student, while counterfactual schools are clustered near zero. The distribution of inputs is more compressed for the faculty-student ratio, which is six faculty per hundred students at MSU, seven at the satellites, and eight at the flagship (Figure 6d). At most colleges attended by the sample, full-time attendance is not the norm (Figure 6e). The treatment shifts students to the top of this distribution.

The socio-economic characteristics of college peers are dramatically different for a student induced to attend the flagship. Nearly 100 percent of flagship parents have graduated from college (Figure 6f). Median parental income at the flagship is over $200,000; the center of the sample distribution is roughly $125,000 (Figure 6g). At most counterfactual colleges, 30 to 50 percent of students receive a Pell Grant; at the flagship only 15 percent do.

VII. Local Average Treatment Effects

The HAIL offer accelerated degree completion and increased degree selectivity. The offer shifted students into higher quality colleges, measured by peer achievement, peer composition, and institutional resources. We now use this experimentally induced variation in college quality to estimate the causal effect of attending a higher quality college on degree completion and projected earnings.

We implement this approach using instrumental variables, with assignment to the HAIL offer as an instrument for college quality. The resulting estimates are average effects for compliers, the students whose college choices shift toward higher quality institutions because of the offer.

VII.A First Stage and Attainment LATEs

We define our endogenous regressor as attending the UM flagship within two years of high school. Overall, the HAIL offer increased the probability of enrolling at the flagship within two years by 18 pp, with an F-statistic of 84.7 (Table 4). This first stage captures the effect of HAIL

[16] The treatment-group bump at about 0.5 is the average of attending a UM satellite and the UM flagship. This average satellite-flagship value is visible on each of the distributions.

on flagship enrollment immediately after high school (15 pp) plus the effect of the deferred admission offer on enrollment two years after high school (three pp).

Attending the flagship increases by 23 pp the likelihood of earning a BA within four years (standard error 11.6 pp, Table 4). This is a large increase relative to the control complier mean of 41 percent. Students induced by HAIL to attend the UM flagship are 65 pp more likely to earn a highly-selective BA. The estimate is significant, with a standard error of 6.3 pp. This is an extremely large effect relative to the control-complier mean of 3.8 percent. Attending the UM flagship leads to a *seventeen-fold increase* in the likelihood of a low-income, high-achieving student earning a highly- or most-selective BA.

VII.B. Projected Earnings Effects

The selectivity of a college degree is strongly predictive of adult earnings. Because our cohorts are still young, we cannot directly observe their earnings. Tax data show that past cohorts who attended the UM flagship earn $78,000 at age 30, compared to $66,000 for those attending Michigan State University and $34,000 for Michigan's two-year colleges (Chetty et al., 2022). As described in the Data section, we use these tax data to project earnings for our sample.

The ITT effect on projected earnings is $3,200 (Table 4). Attending the UM flagship increases earnings of compliers by $17,800 (standard error of $4,300). This is a very large effect (31 percent) relative to the control complier mean of $57,000. In national data, this boost is equivalent to moving from the 46th to the 63rd percentile of the earnings distribution for those with a BA.[17]

VII.C. Plausibility of Exclusion Restriction

Our exclusion restriction is that the HAIL offer affects attainment and earnings only by shifting students into the UM flagship and thereby increasing college quality. That is, the LATE estimates assume that the HAIL offer only affects attainment and earnings for compliers who were induced into the flagship. This assumption allows us to attribute the estimated ITT effects to the 18 percent of students induced into the flagship within two years of high school (Table 4).

This assumption fails if the HAIL offer also affects the outcomes of always-takers. The primary threat is through the effect of certainty of financial aid on time-to-degree. Always-takers in the treatment group received an up-front, four-year guarantee of tuition and fees at the flagship. Always-takers in the control group, by contrast, had no such guarantee, though they ended up getting the same grant aid (Appendix Table 2). We argue in Burland et al. (2023) that it is the

[17] These calculated percentiles are based on the earnings of those aged 28-32 with at least a bachelor's degree in the March Current Population Survey in the years 2014 through 2017. All values are converted to $2025.

certainty in aid, not its generosity, that induced students to apply to the flagship. We now argue that this certainty had no effect on downstream outcomes except through attending the flagship.

Low-income high school students and their parents under-estimate aid and over-estimate tuition (Avery & Kane, 2004; Horn et al, 2003). The HAIL tuition guarantee provided financial certainty for students and families who were nervous about the affordability of a selective college. This altered their applications and ultimately their enrollment.

Our exclusion restriction assumes that the effect of HAIL ended with enrollment. This would hold if students learn about the generosity of flagship aid from either HAIL or from the standard flagship aid offer. The flagship meets the financial need of all of its students; aid is not conditioned on outstanding academic performance. A low-income student who has arrived on campus and in communication with the financial aid office could plausibly have developed the same sense of certainty about aid as a student who gained it through the HAIL shock. In this scenario, the LATE estimates capture the effect of college quality on educational attainment.

Alternatively, our assumption may fail and the HAIL offer may affect the time-to-degree of always-takers in the treatment group. The guarantee may have reduced financial stress or shifted students' planning horizons toward on-time completion. This could raise completion rates, holding fixed the college attended.

We cannot separately identify effects for compliers and always-takers. Under the plausible assumption that any always-taker response is positive, our LATE estimates are biased upward because they attribute always-taker behavioral changes to compliers. The ITT estimates do not rely on this exclusion restriction and remain valid as average effects of being offered HAIL.

VIII. Conclusion

We analyze a randomized intervention that generates plausibly exogenous shifts in college quality. The results provide experimental evidence on how college quality shapes degree trajectories and longer-run economic prospects for disadvantaged students who are qualified to attend the nation's most selective schools. We find large gains in degree selectivity that predict substantial increases in lifetime income and income mobility. Students in the treatment group are roughly twice as likely as those in the control to earn a BA at a highly- or most-selective college, the two highest Barron's rankings.

We summarize the effects of the intervention on BA selectivity in Figure 7. We include for context the values for high-achieving, upper-income students from the same Michigan cohorts.[18]

[18] We define "upper-income" as those not eligible for free- or reduced-price lunch.

The HAIL intervention more than closed the income gap at the top of this selectivity distribution. Among upper-income students, 21 percent earn a BA from a highly- or most-selective college. For the low-income students in our treatment group, the rate is 26 percent. In the control the rate is 14 percent. Most of this increased density at the top is accounted for by a reduction in BAs earned from the three lowest Barron's categories (-9 pp): the rate is 31 percent in the control group and 22 percent in the treatment group.

We find no support for the undermatch hypothesis, which predicts that disadvantaged students will have worse outcomes in a more competitive academic environment. Context and population matter when assessing the threat of undermatch. The students in our sample are well prepared, scoring considerably above their college classmates on admissions tests. They are not academically marginal admits. For this population, shifting to the most competitive university in the state improves academic match while also increasing schooling inputs.

LATE estimates indicate that enrolling at the flagship increases by 65 percentage points the likelihood of earning a BA from a college as selective as the UM flagship. This is extremely large relative to the control-complier mean of four percent. For compliers induced to attend the state flagship, projected earnings are boosted by more than 30 percent.

Attending the flagship also speeds completion: treated compliers are 23 percentage points more likely to earn a BA within four years than control compliers, whose on-time completion rate is 41 percent. At six years, the treatment and control groups are equally likely to have earned a BA. The speedier time to degree will likely put students in the labor force more quickly, increasing lifetime earnings. We plan to study the labor market effects of HAIL as the affected cohorts mature.

References

Angrist, J., Autor, D., & Pallais, A. (2022). Marginal effects of merit aid for low-income students. *The Quarterly Journal of Economics*, *137*(2), 1039–1090. https://doi.org/10.1093/qje/qjab050

Arcidiacono, P., Aucejo, E., Coate, P., & Hotz, V. J. (2014). Affirmative action and university fit: Evidence from Proposition 209. *IZA Journal of Labor Economics*, *3*(1), 7. https://doi.org/10.1186/2193-8997-3-7

Avery, C. & Kane, T.J. (2004) Student perceptions of college opportunities: The Boston COACH Program. In C. M. Hoxby (ed.), *College choices: The economics of where to go, when to go, and how to pay for it,* 355–94, Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226355375.003.0009

Barr, A., & Castleman, B. (2025). Increasing degree attainment among low-income students: The role of intensive advising and college quality. *American Economic Review*, *115*(11), 4075–4103. https://doi.org/10.1257/aer.20240669

Black, D. A., & Smith, J. A. (2006). Estimating the Returns to College Quality with Multiple Proxies for Quality. *Journal of Labor Economics*, *24*(3), 701–728. https://doi.org/10.1086/505067

Bleemer, Z. (2024). Top percent policies and the return to postsecondary selectivity [unpublished manuscript]. https://zacharybleemer.com/wp-content/uploads/2020/10/ELC_Paper.pdf (accessed December 2025).

Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after California's Proposition 209. *The Quarterly Journal of Economics*, *137*(1), 115–160. https://doi.org/10.1093/qje/qjab027

Bound, J., Lovenheim, M. F., & Turner, S. (2012). Increasing time to baccalaureate degree in the United States. *Education Finance and Policy*, *7*(4), 375–424. https://doi.org/10.1162/EDFP_a_00074

Burland, E., Dynarski, S., Michelmore, K., Owen, S., & Raghuraman, S. (2023). The power of certainty: Experimental evidence on the effective design of free tuition programs. *American Economic Review: Insights*, *5*(3), 293–310. https://doi.org/10.1257/aeri.20220094

Burland, E., Camo-Biogradlija, J., Ferrara-Gerson, K., Fields, X., Michelmore, K., Sotherland, N., Stange, K., Thompson, M., and Tompkins-Stange, M. (2025). Bureaucracy and burden: Understanding take-up of a need-based financial aid program. (EdWorkingPaper: 25-1283). Annenberg Institute at Brown University. https://doi.org/10.26300/0ezs-5g46

Chetty, R., Deming, D. J., & Friedman, J. N. (2026). Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges. *The Quarterly Journal of Economics*, 141(1), 51-145. https://doi.org/10.1093/qje/qjaf050

Chetty, R., Friedman, J., Saez, E., Turner, N., & Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the United States. *The Quarterly Journal of Economics,* 135 (3), 1567–633. https://doi.org/10.1093/qje/qjaa005.

Chetty, R., Friedman, J., Saez, E., Turner, N., & Yagan, D. (2022). *Replication data for: Income segregation and intergenerational mobility across colleges in the United States* (Version 1.0) [Dataset]. Harvard Dataverse. https://doi.org/10.7910/DVN/RYLJKZ

Cohodes, S., Ho, H., & Robles, S. (2024). Stem summer programs for underrepresented youth increase stem degrees (No. w30227). *National Bureau of Economic Research*. https://doi.org/10.3386/w30227

Cohodes, S. R., & Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, *6*(4), 251–285. https://doi.org/10.1257/app.6.4.251

College Scorecard. (2015). *College Scorecard Institution-Level Data* [Dataset]. https://collegescorecard.ed.gov/data/ (accessed December 17, 2025)

College Scorecard. (2023). *College Scorecard Institution-Level Data* [Dataset]. https://collegescorecard.ed.gov/data/ (accessed December 17, 2025)

Dale, S. B., & Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, *117*(4), 1491–1527. https://doi.org/10.1162/003355302320935089

Dale, S. B. & Krueger, A. B., (2014). Estimating the effects of college characteristics over the career using administrative data. *Journal of Human Resources, 49*(2), pp. 323-358. https://doi-org/10.3368/jhr.49.2.323

Dillon, E. W., & Smith, J. A. (2020). The Consequences of academic match between students and colleges. *Journal of Human Resources*, *55*(3), 767–808. https://doi.org/10.3368/jhr.55.3.0818-9702R1

Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721–1756. https://doi.org/10.1257/aer.20200451

Goldrick-Rab, S. (2016). *Paying the price: College costs, financial aid, and the betrayal of the American dream* (D. M. Anderson & P. Kinsley, Eds.). The University of Chicago Press.

Goodman, J., Hurwitz, M., & Smith, J. (2017). Access to 4-Year public colleges and degree completion. *Journal of Labor Economics*, *35*(3), 829–867. https://doi.org/10.1086/690818

Hoekstra, M. (2009). The Effect of attending the flagship state university on earnings: A discontinuity-based approach. *Review of Economics and Statistics*, *91*(4), 717–724. https://doi.org/10.1162/rest.91.4.717

Horn, L. J., Xianglei C., & Chapman, C. (2003). Getting ready to pay for college: What students and their parents know about the cost of college tuition and what they are doing to find out. U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubs2003/2003030.pdf

Hoxby, C., & Avery, C. (2012). The missing "one-offs": The hidden supply of high-achieving, low-income students (No. w18586). National Bureau of Economic Research. https//doi.org/10.3386/w18586

Hyman, J., 2017. ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice. *Education Finance and Policy,* 12(3), pp.281-311. https://doi.org/10.1162/EDFP_a_00206

Jack, A. A. (2019). *The privileged poor: How elite colleges are failing disadvantaged students*. Harvard University Press. https://doi.org/10.4159/9780674239647

Jack, A. A., & Black, Z. (2024). Belonging and boundaries at an elite university. *Social Problems*, *71*(4), 996–1013. https://doi.org/10.1093/socpro/spac051

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119. https://doi.org/10.1111/j.1468-0262.2007.00733.x

Kozakowski, Whitney. (2023). *Are four-year public colleges engines for economic mobility? Evidence from statewide admissions thresholds*. (EdWorkingPaper: 23-727). Annenberg Institute at Brown University. https://doi.org/10.26300/VAPT-4E25

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. The Review of Economic Studies, 76(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Light, A., & Strayer, W. (2000). Determinants of college completion: School quality or student ability? *The Journal of Human Resources*, *35*(2), 299. https://doi.org/10.2307/146327

Lovenheim, M., & Smith, J. (2023). Returns to different postsecondary investments: Institution type, academic programs, and credentials. In Hanushek, E. A., Machin, S., Woessmann, L. , *Handbook of the Economics of Education* (Vol. 6, pp. 187–318). Elsevier. https://doi.org/10.1016/bs.hesedu.2022.11.006

Miller, L. (2025). Switching schools: Effects of college transfers. (EdWorkingPaper: 25-1159). Annenberg Institute at Brown University. https://doi.org/10.26300/4khw-c054

Mountjoy, J. (2026). Marginal returns to public universities. *The Quarterly Journal of Economics*, 141(1): Pages 429–497. https://doi.org/10.1093/qje/qjaf055

National Academies of Sciences, Engineering, and Medicine. (2025). Postsecondary Education. In *Economic and Social Mobility: New Directions for Data, Research, and Policy*. National Academies Press. https://doi.org/10.17226/28456.

National Science Board. (2024). *The State of U.S. Science and Engineering 2024*. National Science Foundation. https://ncses.nsf.gov/pubs/nsb20243 (accessed December 2025).

U.S. Department of Education, National Center for Education Statistics. (n.d.). "College Navigator." College Navigator. https://nces.ed.gov/collegenavigator/ (accessed December 2025).

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS). (2024). *Digest of Education Statistics, Table 326.10.* National Center for Education Statistics. https://nces.ed.gov/programs/digest/d24/tables/dt24_326.10.asp.

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS). (2016). [College IPEDS enrollment-full-time-equivalent; College IPEDS fall-enrollment race sex, Education Data Portal; College IPEDS finance]. Education Data Portal (Version 0.24.0), Urban Institute, made available under the ODC Attribution License. https://educationdata.urban.org/documentation/ (accessed December 2025).

U.S. Department of Education, National Center for Education Statistics. (2017a). Table 1.2-A: Among 2011–12 first-time postsecondary students who began at a 2-year public college, percentage distribution of 6-year attainment and retention status at any institution attended, by selected personal characteristics: 2012–17. Data Lab Tables Library. Retrieved from https://nces.ed.gov/datalab/table/library/detail/14326.

U.S. Department of Education, National Center for Education Statistics. (2017b). Table 226.40: SAT mean scores, standard deviations, and score ranges for high school seniors, and percentage of the graduating class taking the SAT, by state: 2017. Retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_226.40.asp

U.S. Department of Homeland Security. (2022). *DHS STEM Designated Degree Program List*. https://www.ice.gov/doclib/sevis/pdf/stemList2022.pdf (accessed April 2023).

Zimmerman, S. D. (2014). The Returns to college admission for academically marginal students. *Journal of Labor Economics*, *32*(4), 711–754. https://doi.org/10.1086/676661

Table 1
High-School-Level Summary Statistics by Treatment Arm

Characteristic	Control mean	Treated mean	P-value
Southeast MI school	0.401	0.394	0.864
	(0.029)	(0.030)	
West, Central, Northern MI school	0.599	0.606	0.864
	(0.029)	(0.030)	
Suburban school	0.340	0.360	0.592
	(0.028)	(0.029)	
City school	0.129	0.100	0.273
	(0.020)	(0.018)	
Town/rural school	0.530	0.540	0.838
	(0.030)	(0.030)	
Distance of school from UM (miles)	93.2	96.4	0.597
	(4.691)	(4.824)	
Average number of sample students	3.8	3.9	0.613
	(0.177)	(0.210)	
Proportion eligible for free lunch	0.709	0.692	0.450
	(0.015)	(0.015)	
Proportion eligible for reduced-price lunch	0.291	0.308	0.450
	(0.015)	(0.015)	
Proportion female	0.571	0.605	0.102
	(0.016)	(0.015)	
Proportion male	0.429	0.395	0.102
	(0.016)	(0.015)	
Proportion Black	0.094	0.087	0.685
	(0.013)	(0.011)	
Proportion Hispanic	0.053	0.057	0.681
	(0.008)	(0.009)	
Proportion Asian American	0.061	0.057	0.718
	(0.008)	(0.009)	
Proportion American Indian, Alaska Native, or Native Hawaiian	0.019	0.012	0.253
	(0.006)	(0.004)	
Proportion White	0.772	0.787	0.566
	(0.017)	(0.016)	
Average SAT	1254	1259	0.200
	(2.805)	(3.017)	
Average GPA	3.82	3.83	0.263
	(0.006)	(0.006)	
Proportion sent SAT scores to UM	0.365	0.377	0.580
	(0.015)	(0.015)	
F-test p-value		0.477	
Number of school-years	526	500	1,026
Number of students	1,978	1,932	3,910

Notes: All analyses are conducted at the high school-by-cohort level. p-values are from a t-test of the coefficient on treatment status from a regression of the characteristic on treatment and strata dummies. F-test is from a joint significance test predicting treatment based on the characteristics listed here as well as strata dummies. Standard errors clustered at the high-school-level in parentheses.

Table 2
Estimated Effect of HAIL on Degree Attainment and Degree Type

Outcome	Effect of HAIL offer (ITT)	Control mean
Enroll UM Flagship within 1 year	0.151	0.118
	(0.019)	
Enroll UM Flagship within 2 years	0.178	0.124
	(0.019)	
BA attainment		
Within 4 years	0.041	0.391
	(0.022)	
Within 6 years	-0.006	0.625
	(0.021)	
Highly competitive+ BA attainment		
Within 4 years	0.113	0.103
	(0.017)	
Within 6 years	0.117	0.144
	(0.019)	
STEM BA		
Within 4 years	0.018	0.149
	(0.016)	
Within 6 years	-0.010	0.257
	(0.019)	
N school-cohorts	1,026	

Notes: All analyses are conducted at the high-school-by-cohort-level. Robust standard errors reported in parentheses. Treatment effect coefficients are estimates of β_1 from Equation (1). BA attainment is measured by the end of the fourth and sixth summers following high school, and is not conditional on initial enrollment.

Table 3
Estimated Effect of HAIL on Transfer Pathways

	Transfer in Fall 2		Transfer in Fall 3	
Outcome	ITT: Effect of HAIL Offer	Control mean	ITT: Effect of HAIL Offer	Control mean
Any Transfer	0.023	0.080	-0.021	0.070
	(0.012)		(0.011)	
Institution type				
2-year to 4-year	-0.007	0.020	-0.014	0.031
	(0.005)		(0.008)	
4-year to 2-year	-0.002	0.027	-0.004	0.011
	(0.007)		(0.004)	
2-year to 2-year	0.005	0.002	-0.002	0.003
	(0.003)		(0.002)	
4-year to 4-year	0.028	0.031	-0.001	0.023
	(0.009)		(0.006)	
Transfer into UM Flagship				
From 2-year or non-UM 4-year	-0.002	0.004	-0.003	0.004
	(0.002)		(0.002)	
From UM Satellite	0.029	0.001	0.001	0.001
	(0.005)		(0.001)	
Transfer from UM Flagship				
To a 2-year or non-UM 4-year	0.010	0.000	0.003	0.002
	(0.004)		(0.003)	
To a UM Satellite	0.003	0.000	0.001	0.000
	(0.002)		(0.001)	
N school-cohorts	1026			

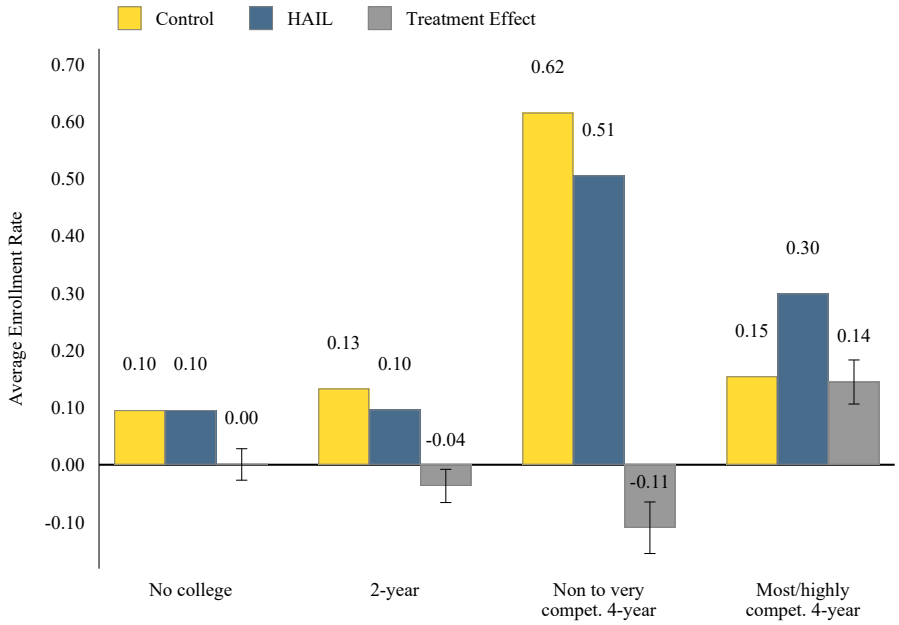
Notes: All analyses are conducted at the high-school-by-cohort-level. Robust standard errors reported in parentheses. Treatment effect coefficients are estimates of β_1 from Equation (1). Transfers are defined as having different enrollment from one fall to the following fall (fall 1 to fall 2, fall 2 to fall 3).

Table 4
Local Average Treatment Effects of University of Michigan Enrollment on Bachelor's Degree Attainment and Projected Earnings

	ITT		LATE	
	Effect of HAIL offer	Control mean	Effect of UM attendance	Control complier mean
First Stage				
Enroll UM Flagship within 2 years	0.178	0.124		
	(0.019)			
F-statistic	84.702			
ITT and LATE				
BA attainment				
Within 4 years	0.041	0.391	0.230	0.406
	(0.022)		(0.116)	
Within 6 years	-0.006	0.625	-0.036	0.791
	(0.021)		(0.121)	
Highly Competitive+ BA attainment				
Within 6 years	0.117	0.144	0.658	0.038
	(0.019)		(0.063)	
STEM BA				
Within 6 years	0.018	0.149	0.102	0.224
	(0.016)		(0.085)	
Projected Earnings	3,163	53,154	17,771	56,544
Projected Earnings at Age 30	(933)		(4,301)	
N school-cohorts	1,026			

Notes: All analyses are conducted at the high-school-by-cohort-level. Robust standard errors reported in parentheses. ITT coefficients are estimates of β_1 in Equation (1). LATEs are estimates of γ_1 in Equation (3), estimated using 2SLS. BA attainment is measured by the end of the fourth and sixth summers following high school, and is not conditional on initial enrollment. See Appendix B for details on how STEM is calculated. Median earnings are from Opportunity Insights, measured in 2025 values, adjusting for inflation using the consumer price index (CPI-U). Students are assigned the median age 30 earnings for their modal undergraduate institution attended within 4 years of high school. Students with two modal institutions are assigned the average of the age 30 median income for the two modal institutions. Control complier means are calculated at the student level, inverse weighted by size of high school cohort.

Panel A. First-year college enrollment by selectivity category



Panel B. First-year college enrollment at University of Michigan, by campus

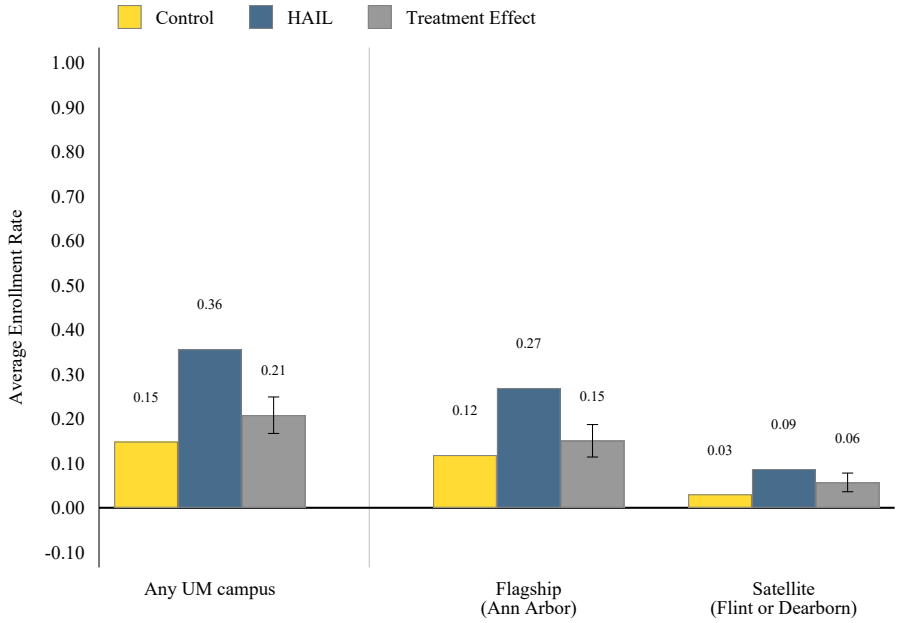
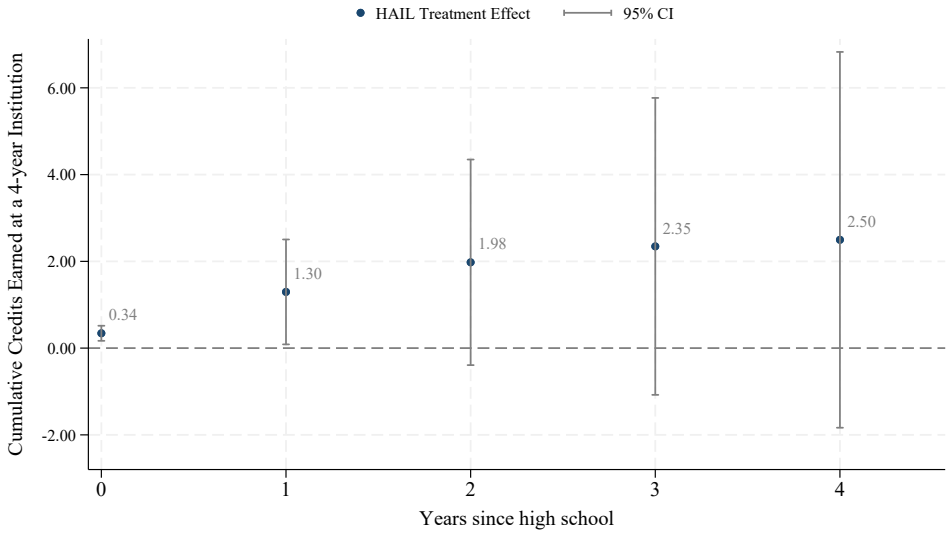


Figure 1
Estimated Effect of HAIL on First-Year College Enrollment

Notes: Figure plots the rate of enrollment in the first year following high school by Barron's institution competitiveness, and by UM flagship (Ann Arbor) and UM Satellites (Flint or Dearborn). All analyses are conducted at the high-school-by-cohort-level. Treatment effect coefficients are estimates of β_1 in Equation 1. Ninety-five percent confidence intervals on treatment effects are based on robust standard errors.

Panel A. Cumulative Credits Earned



Panel B. On-track for 4-year BA Completion

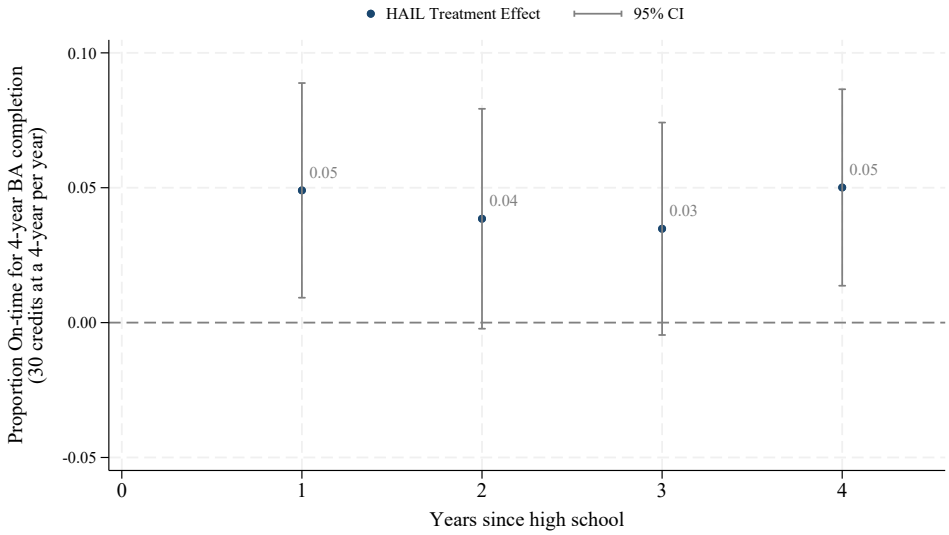
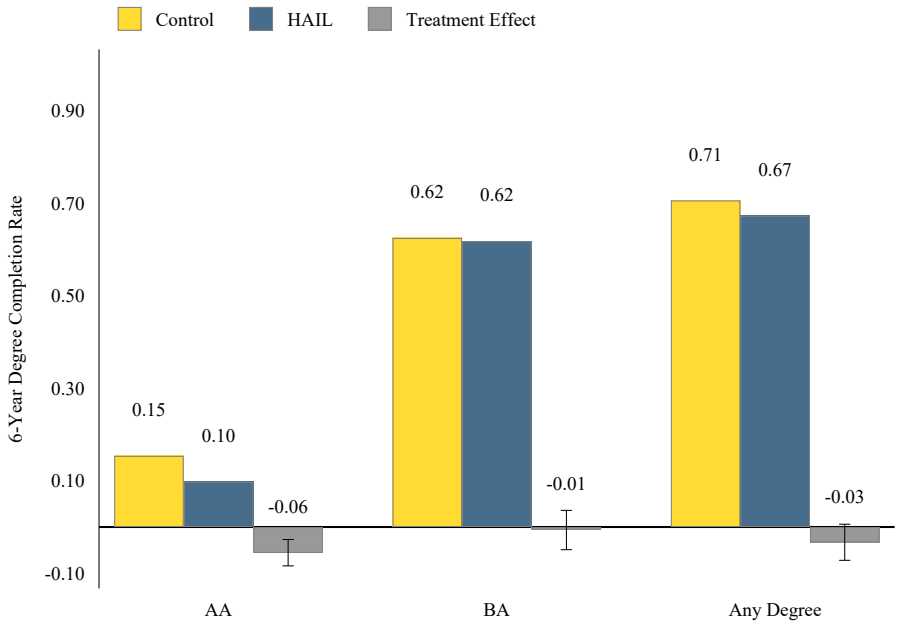


Figure 2
Effect of HAIL on Credit Accumulation at 4-year Institutions

Notes: All analyses are conducted at the high-school-by-cohort-level. Robust standard errors are reported in parentheses. Treatment effect coefficients are estimates of β_1 in Equation 1. Credits are calculated at the term-level then aggregated to the student-year level. Measured credits include undergraduate coursework taken at a given institution. Annual credits are measured at the end of the summer term and include fall, spring, and summer credits. Year 0 includes the summer between high school and fall 1. Credits are imputed for students who have an undergraduate enrollment record in NSC but are missing credit information in STARR using a regression of the credit outcome on selectivity, enrollment-status, years since high school, and semester-type (fall, spring, summer) for control students with non-missing credits, then summed by year.

Panel A. Any degree attainment



Panel B. Degree combinations

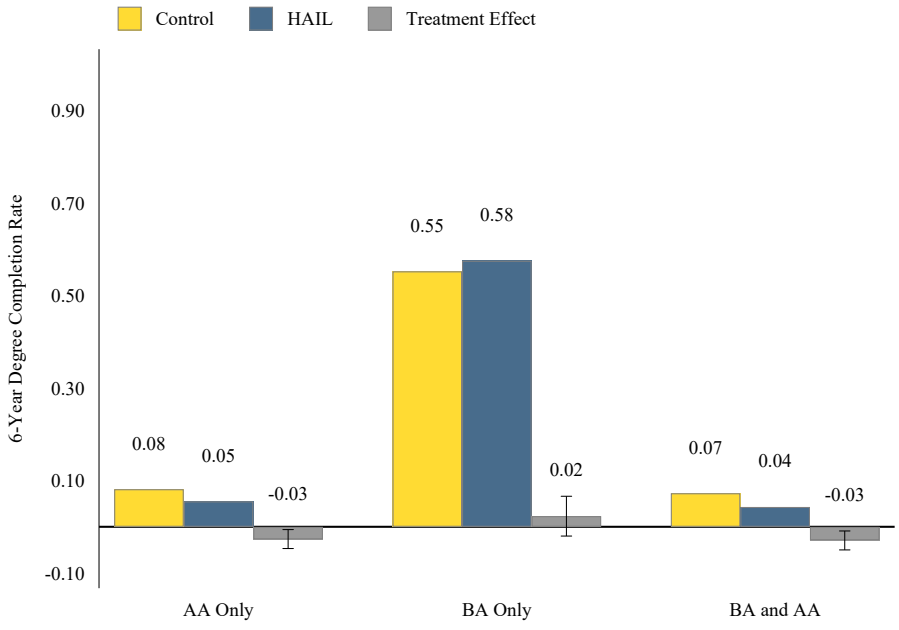


Figure 3
Estimated Effect of HAIL on Degree Attainment

Notes: Figure plots the rate of degree attainment within 6 years of high school. All analyses are conducted at the high-school-by-cohort-level. Treatment effect coefficients are estimates of β_1 in Equation 1. Ninety-five percent confidence intervals on treatment effects are based on robust standard errors.

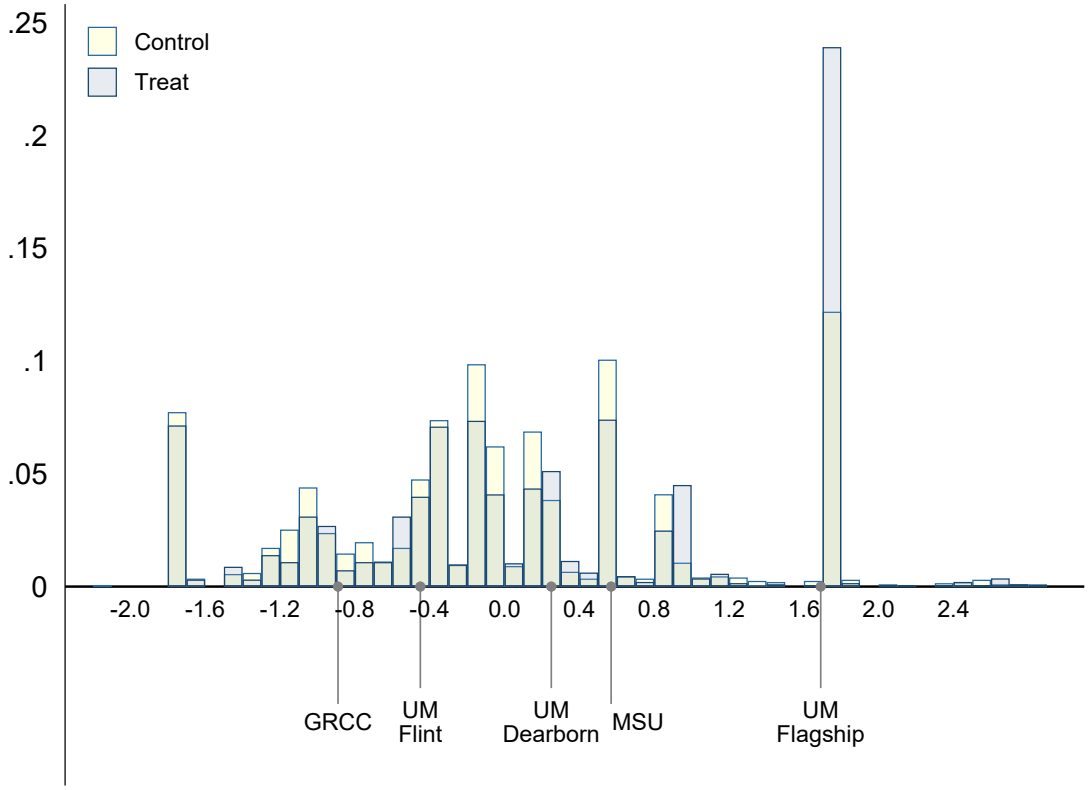
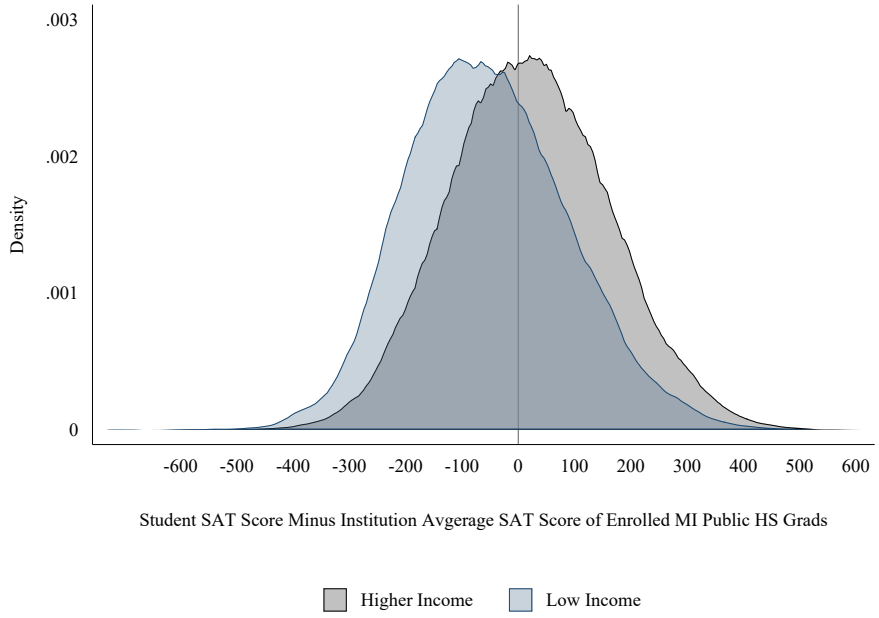


Figure 4
Distribution of institution mean SAT scores at enrolled institution

Notes: Distribution is at the student-level. Institution mean SAT is calculated using the test scores of Michigan public high school who finish high school 2013-2014 and attend a given institution. For consistency, we convert all ACT scores to SAT scores using ACT-SAT concordance tables provided by the College Board. For institutions with fewer than 30 Michigan public high school students enrolled, the mean score for the institution tier is used. Students not enrolled are assigned the mean score for students who do not enroll in the first fall. Students are assigned an average of the institution mean SAT for the 1st institution enrolled in Fall 1 and the 1st institution enrolled in Fall 2 (to get a weighted average "college quality" for years 1-2).

Panel A. All Michigan students



Panel B. HAIL treated and control students

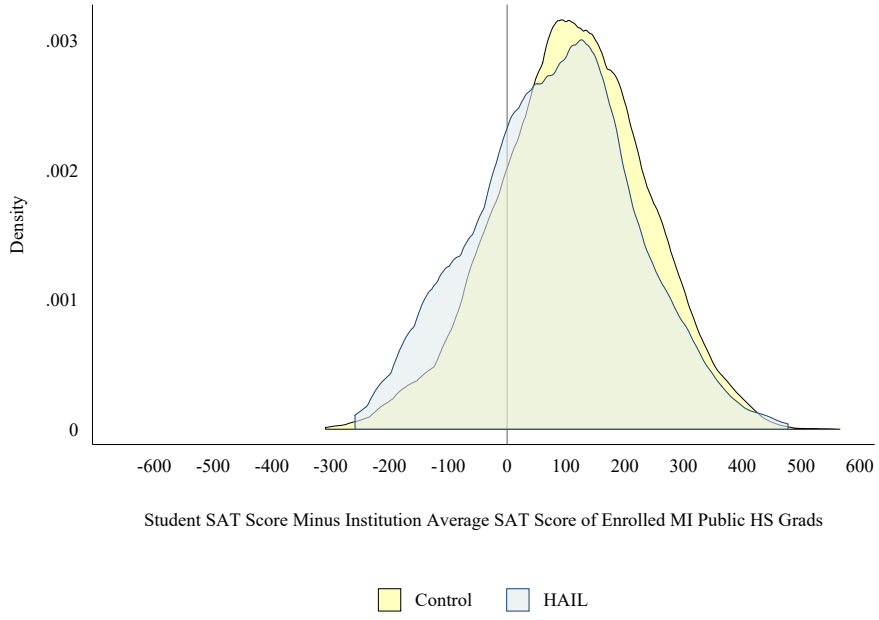
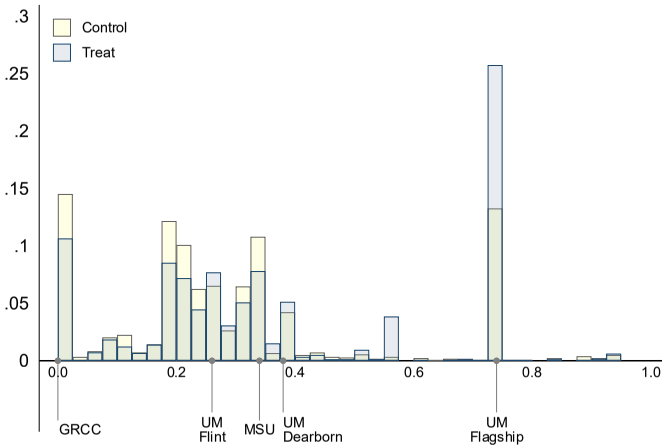
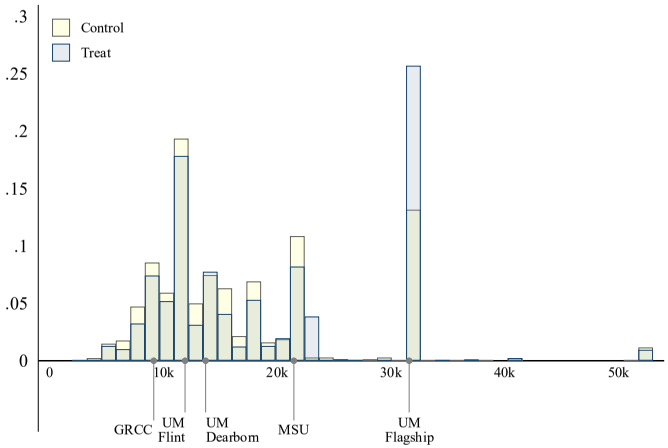


Figure 5
Average SAT Scores Relative to the Average SAT Score of Institutional Peers

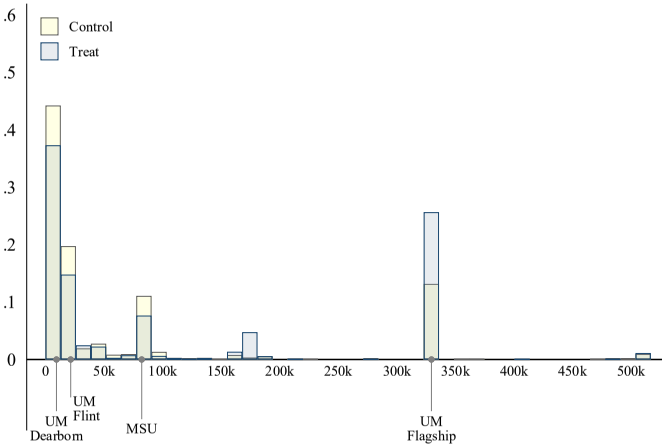
Notes: Figures plot the difference between a student's SAT score and the average SAT score of their peers at the institution they attend. Distributions are at the student-level. Institution mean SAT is calculated using the test scores of Michigan public school students finishing high school in 2013-2014 who attend a given institution in the first fall following high school. For consistency, we convert all ACT scores to SAT scores using ACT-SAT concordance tables provided by the College Board. For institutions with fewer than 30 Michigan public high school students enrolled, the average score for the institution tier is used. Students not enrolled are assigned the average score for students who do not enroll in the first fall. Panel A includes all Michigan public high school students expected to graduate in the same cohorts as the HAIL sample (2016-2017), not restricted by academic achievement.



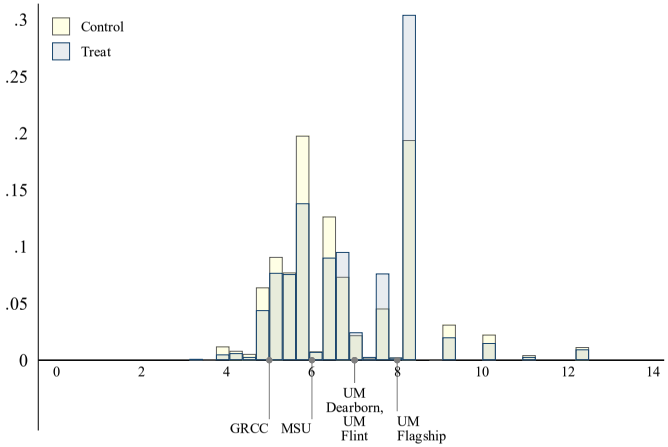
a) Rejection rate (1-acceptance rate)



(b) Instructional expenditures per FTE student

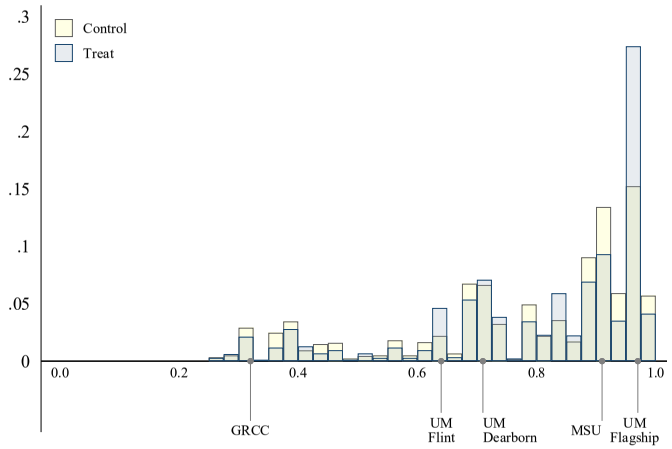


(c) Endowment per FTE student

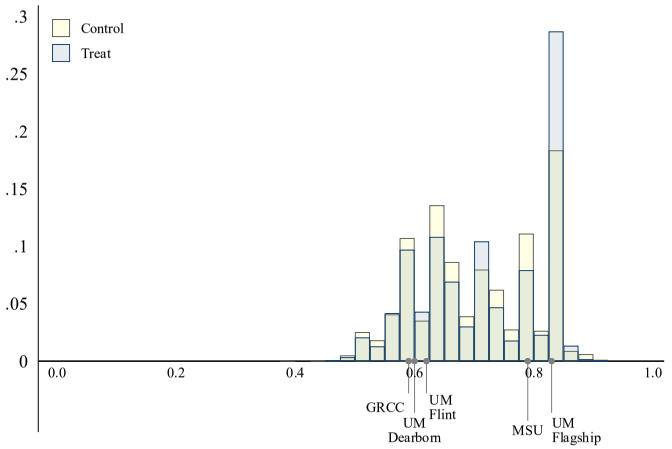


(d) Instructional faculty per 100 students

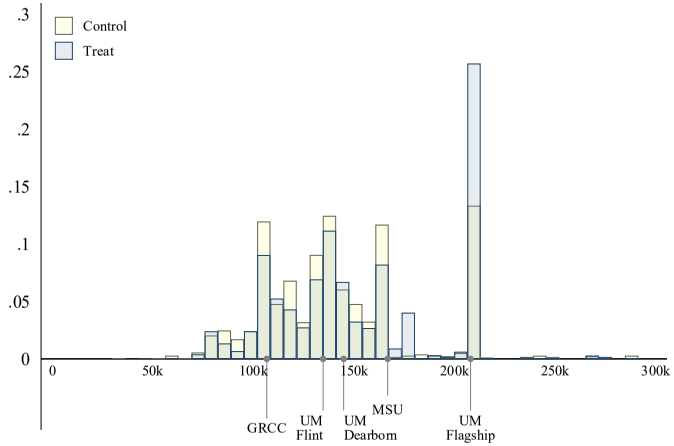
Figure 6
Distribution of Peer and Institutional Characteristics



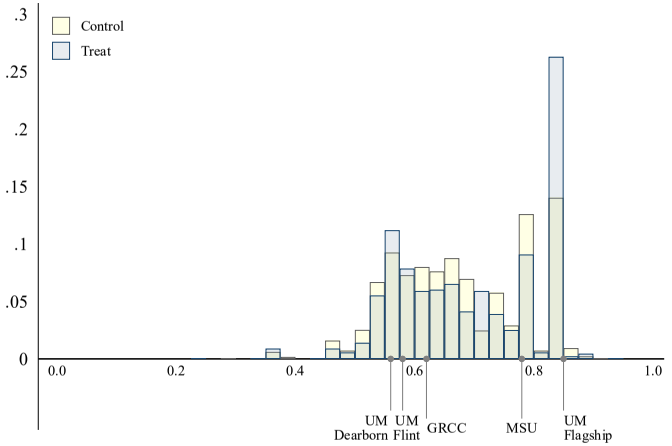
(e) Share enrolled full time



(f) Share with a parent who graduated from college



(g) Median parental income



(h) Share not receiving a Pell Grant

Figure 6
Distribution of Peer and Institutional Characteristics (cont.)

Notes: Distributions are at the student-level. We use Michigan State University and Grand Rapids Community College as illustrative examples here because they are the most common 4-year and 2-year institutions attended by students in the control group. Individuals are assigned an average of the institution characteristics for their year 1 institution (first institution attended) and their year 2 institution (first institution attended). For students who only enrolled in year 1, they are assigned the values for their year 1 institution. The same logic is applied for students who only enroll in year 2. Students who do not enroll in years 1 or 2 are dropped. Institutional characteristics are based on data from the 2015-16 academic year from IPEDS and College Scorecard. Parental Income is from Opportunity Insights, for the 1984 birth cohort. Instructional expenditures, endowment, and parental income are measured in 2025 dollars, adjusting for inflation using the consumer price index (CPI-U). Parent education is based on the College Scorecard definition of first-generation from prior to 2017-2018: "if at least one parent's status was 'completed college' then this was coded as non-first generation." Source: IPEDS, College Scorecard, Opportunity Insights.

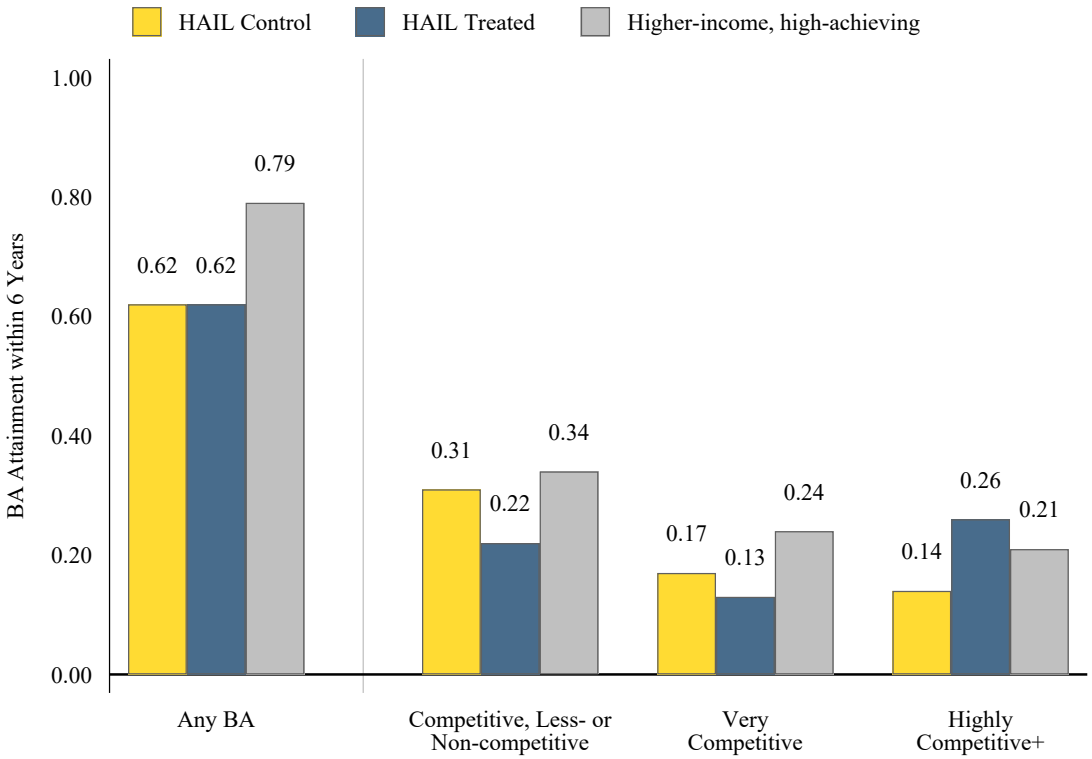


Figure 7
Selectivity of Bachelor's Degrees Earned by High-Achieving Michigan Students, by Income and HAIL Treatment Status

Notes: Sample is eleventh grade students in Michigan public schools in 2015 and 2016 who meet HAIL GPA and ACT/SAT criteria. Higher-income students are those not eligible for free- or reduced-price lunch. Degree selectivity is based on Barron's competitiveness index. Degree attainment rates for HAIL treatment and control groups use our high-school-by-cohort-level measures. Average degree attainment for higher-income students is calculated at the student level.

A Pathway to Mobility: Shifting Low-Income Students to Higher-Quality Degrees
Online Appendix

Elizabeth Burland, Nora Delaney, Susan Dynarski, Katherine Leu, CJ Libassi, Katherine Michelmore, Stephanie Owen, Mary Quiroga, and Elizabeth Salinas

A Appendix Tables and Figures

Appendix Table 1
Summary of Literature on Effects of College Quality

Citation	Average SAT score of sample	T vs C (test scores, SDs, ITT)	4-year BA attainment		6-year BA attainment		Earnings		Context
			Control mean	Treatment Effect (ITT)	Control mean	Treatment Effect (ITT)	Control mean	Treatment effect (%, ITT)	
Present Analysis	**1254**	**0.270**	**0.390**	**0.040**	**0.620**	**-0.010**	**37,907**	**0.072**	**MI**
Chetty et al. (2026)	1374							0.040	National
Mountjoy (2026)	916	0.075	0.100	0.008	0.300	0.018	40,815	0.013	TX
Barr & Castleman (2025)			0.268	0.062	0.528	0.076			MA, NY, IL
Bleemer (2024)	1043	0.220			0.600	0.022	70,000	0.039	CA
Cohodes et al. (2024)	1436		0.532	0.082	0.736	0.025	104,213	0.118	National
Kozakowski (2023)	878	0.060			0.390	0.036	30,577	0.064	MA
Bleemer (2022)+	1025	0.150	0.478	0.009	0.746	0.007	60,888	0.050	CA
Angrist et al. (2022)	1100		0.320	-0.028	0.636	0.081			NE
Goodman et al. (2017)	906	0.020			0.370	0.009			GA
Dale & Krueger (2014)									
Overall	1210						139,698	-0.008	National
Black and Hispanic	1024						98,472	0.076	National
Cohodes & Goodman (2014)+	1279	0.018	0.462	0.019	0.631	0.025			MA
Zimmerman (2014)	943		0.050	-0.007	0.200	0.057	28,963	0.051	FL
Hoekstra (2009)	1000							0.135	Unknown

Notes: Notes: + signifies the effects are reverse coded to align with the direction of the other results for easier comparison. Notes on average SAT score of the sample: We convert Angrist et al. (2022) reported ACT scores to SAT scores. For estimates from Cohodes and Goodman (2014) and Cohodes, Ho, and Robles (2023), we convert reported standardized test scores (in SD units) to SAT scores using NCES-reported SAT means and SDs. Cohodes and Goodman (2014) report that those in the RD sample score about 0.90 standard deviation above average MCAS scores relative to all high school grads in Massachusetts, we use Massachusetts specific estimates to convert this to an equivalent SAT score. Cohodes, Ho, and Robles (2023) report a standardized math score (a mix of PSAT, SAT, and ACT scores), an average of 1.929 sd which we convert to an equivalent SAT score using the national mean and standard deviation from 2017. For the paper from Hoekstra, the SAT score we report is based on this description in the text: "the average SAT scores of the entering classes at the flagship over the period of time studied were between 1000 and 1100" (2009, p. 717). We use the lower number since that range is the average, and the RD is for the marginal student.The T vs C contrast column represents the average test scores of students attending institutions in the treated group relative to the comparison group. Unless otherwise noted, these reflect differences in SAT scores between treatment and comparison groups. Goodman et al. (2017) report PSAT scores rather than SAT scores, Mountjoy (2025) reports 11th grade standardized test scores, and Kozakowski reports MCAS standardized test scores from Massachusetts. Cohodes and Goodman (2014) only report SAT math scores, not composite. All other papers use SAT composite test scores. SAT score treatment effects converted into standard deviation units using our institution distribution, where a standard deviation unit is approximately 133 points.

Appendix Table 2
Estimated Effect of HAIL on Financial Aid Receipt at the UM Flagship
Adapted from Burland et al. (2023, p.304)

	Panel A: First Year			Panel B: Sum of Years 1-4		
	Control Mean	HAIL mean	P-value	Control Mean	HAIL mean	P-value
Grants	$24,729	$24,207	0.360	$106,643	$105,735	0.720
	(497)	(235)		(2230)	(1011)	
Loans	$1,339	$1,766	0.125	$5,293	$7,282	0.042
	(202)	(164)		(772)	(592)	
Proportion with Grants >= Tuition	0.932	1.000	0.000	0.896	0.997	0.000
	(0.017)	(0.000)		(0.025)	(0.003)	
Expected Family Contribution	$1,902	$2,078	0.531	$8,264	$8,812	0.601
	(269)	(206)		(1220)	(827)	
Cost of Tuition	$14,672			$66,199		
	(403)			(398)		
Number of students	237	465		193	347	
Number of students in the study	1,978	1,932		1,978	1,932	

Notes: Adapted from Table 3 in "The power of certainty: Experimental evidence on the effective design of free tuition programs," by Burland, E., Dynarski, S., Michelmore, K., Owen, S. and Raghuraman, S., 2023. American Economic Review: Insights, 5(3): 293–310. (https://doi.org/10.1257/aeri.20220094). Analysis done at the student-level. Includes only students enrolled at the UM Flagship full time for full first year (panel A) or full time for all four years (panel B) and who have financial aid data reported. Standard errors are clustered at the high-school-by-cohort-level. Includes zeros for students who receive no aid. "Grants" includes all institutional and departmental scholarships and grants, federal grants, state grants and scholarships, private scholarships, and other departmental aid. EFC is capped at the cost of attendance, as determined by UM (includes tuition, fees, books, and supplies, room and board, transportation, and personal expenses).

Appendix Table 3
Estimated Effect of HAIL on Cumulative Credits Earned

	Panel A: Cumulative Credits			
	Credits Earned at 4-Year Colleges		Credits Earned at 2-Year Colleges	
	ITT: Effect of HAIL Offer	Control mean	ITT: Effect of HAIL Offer	Control mean
Pre-college summer	0.342	0.245	0.000	0.056
	(0.089)		(0.024)	
After year 4	2.497	78.145	-1.336	8.533
	(2.205)		(0.875)	
After year 6	1.189	85.612	-1.402	9.317
	(2.251)		(0.932)	
N school-cohorts	1026		1026	
	Panel B: Accumulated 30 Credits Per Year			
	Credits Earned at 4-Year Colleges		Credits Earned at 2-Year Colleges	
	ITT: Effect of HAIL Offer	Control mean	ITT: Effect of HAIL Offer	Control mean
30 credits, by end of year 1	0.049	0.306	-0.001	0.029
	(0.020)		(0.007)	
60 credits, by end of year 2	0.038	0.288	0.004	0.011
	(0.021)		(0.005)	
90 credits, by end of year 3	0.035	0.277	0.000	0.000
	(0.020)		(0.000)	
120 credits, by end of year 4	0.050	0.209	0.000	0.000
	(0.019)		(0.000)	
N school-cohorts	1026		1026	

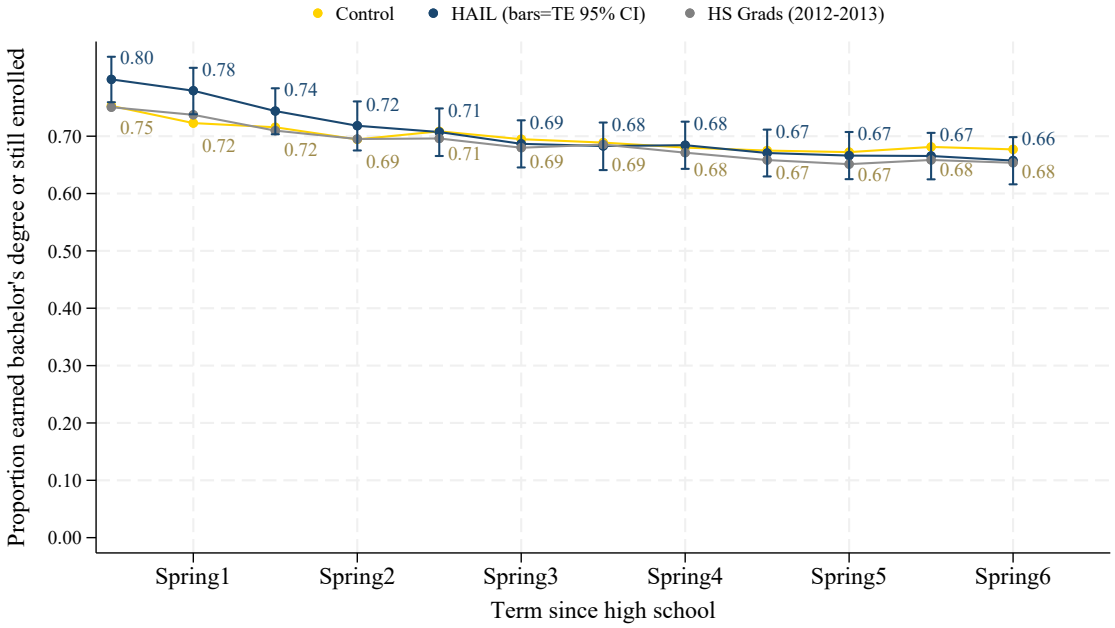
Notes: All analyses done at the high-school-by-cohort-level. Robust standard errors reported in parentheses. Treatment effect coefficients are from estimating Equation (1), the coefficient reports estimates of β_1. Credits calculated at the term-level and include undergraduate coursework taken at a given institution and collapsed to the student-year level. Credits are imputed for students who have an undergraduate enrollment record in NSC but are missing credit information in STARR using a regression of the credit outcome on selectivity, enrollment-status, year since high school, and semester-type (fall, spring, summer) for control students with non-missing credits, then summed by year.

Appendix Table 4
Estimated Effect of HAIL on Institutional characteristics

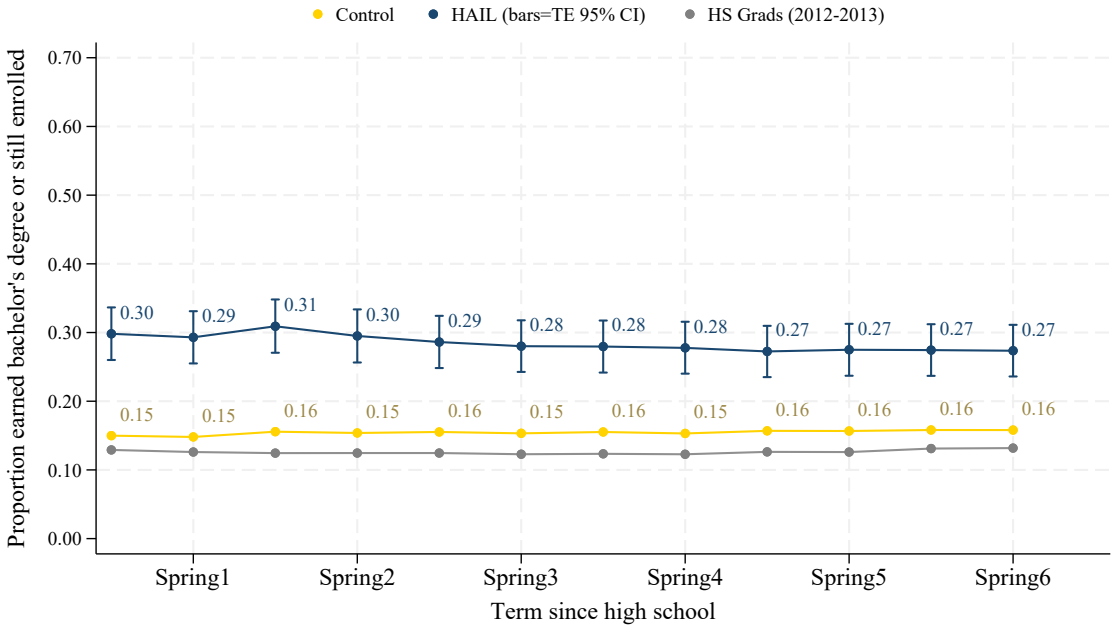
Outcome	Effect of HAIL offer (ITT)	Control mean
Rejection rate (1-acceptance rate)	0.094	0.287
	(0.012)	
Instructional expenditures per FTE student	2,806	16,250
	(444)	
Endowment per FTE student	50,480	72,018
	(6488)	
Instructional faculty per 100 students	0.329	6.668
	(0.072)	
Share enrolled full time	0.030	0.779
	(0.009)	
Median parental income	13,328	140,638
	(2025)	
Share with a parent who graduated from college	0.018	0.703
	(0.005)	
Share not receiving a Pell Grant	0.025	0.675
	(0.006)	
N school-cohorts	998	
N students	3,621	

Notes: All analyses done at the high-school-by-cohort-level. Robust standard errors reported in parentheses. Treatment effect coefficients are from estimating Equation (1), the coefficient reports estimates of β_1. Individuals are assigned an average of the institution characteristics for their year 1 institution (first institution attended) and their year 2 institution (first institution attended). For students who only enrolled in year 1, they are assigned the values for their year 1 institution. The same logic is applied for students who only enroll in year 2. Students who do not enroll in years 1 or 2 are dropped. Parental income, endowment per FTE, and instructional expenditures per FTE are measured in 2025 values, adjusting for inflation using the consumer price index (CPI-U).
Source: median parental income is from Opportunity Insights, and all other variables are from College Scorecard.

Panel A. BA or still enrolled at a four-year institution



Panel B. Highly competitive BA, or still enrolled at a highly competitive institution



Appendix Figure 1
Pandemic Impact on Persistence and Degree Attainment

Notes: Estimated at the high-school-by-cohort-level. March 2020 (start of COVID19) was midway through spring 4 (HAIL cohort 1) and 3 (HAIL cohort 2), potentially impacting enrollments starting in fall 5 or fall 4 (for cohorts 1 and 2, respectively). For reference, we include estimates for pre-HAIL cohorts expected to complete high school in 2013 and 2014, and who meet the economic and academic criteria for the first HAIL cohort. Their persistence/graduation would be unaffected by the COVID-19 pandemic which started 6 to 7 years post high school.

B Outcomes and Measurement

B.1 Enrollment

We use enrollment from NSC, which we structure into three terms per academic year—fall, spring, and summer—using NSC enrollment dates. We measure enrollment from expected high school graduation. The first HAIL cohort graduated high school in the spring of 2016, therefore "fall 1" is fall 2016, "spring 1" is spring 2017, and "summer 1" is summer 2017. In each term, we create mutually-exclusive enrollment outcomes based on the first institution enrolled in a given term. In the transcript data, we enforce the same three term structure that aligns with the NSC enrollment outcomes, which allows us to merge these data at the student-by-term level.

Here, we only capture undergraduate enrollment, which we define as any enrollment terms that occur prior to a student earning their first bachelor's degree in NSC. We also exclude terms defined as "Graduate" in the student transcript data. We also exclude terms where a student was enrolled in a non-degree program according to the student transcript data, or terms where a student was enrolled for fewer than 15 days, which corresponds with the timing to withdraw and still receive a refund on tuition paid.

B.2 Degree Attainment

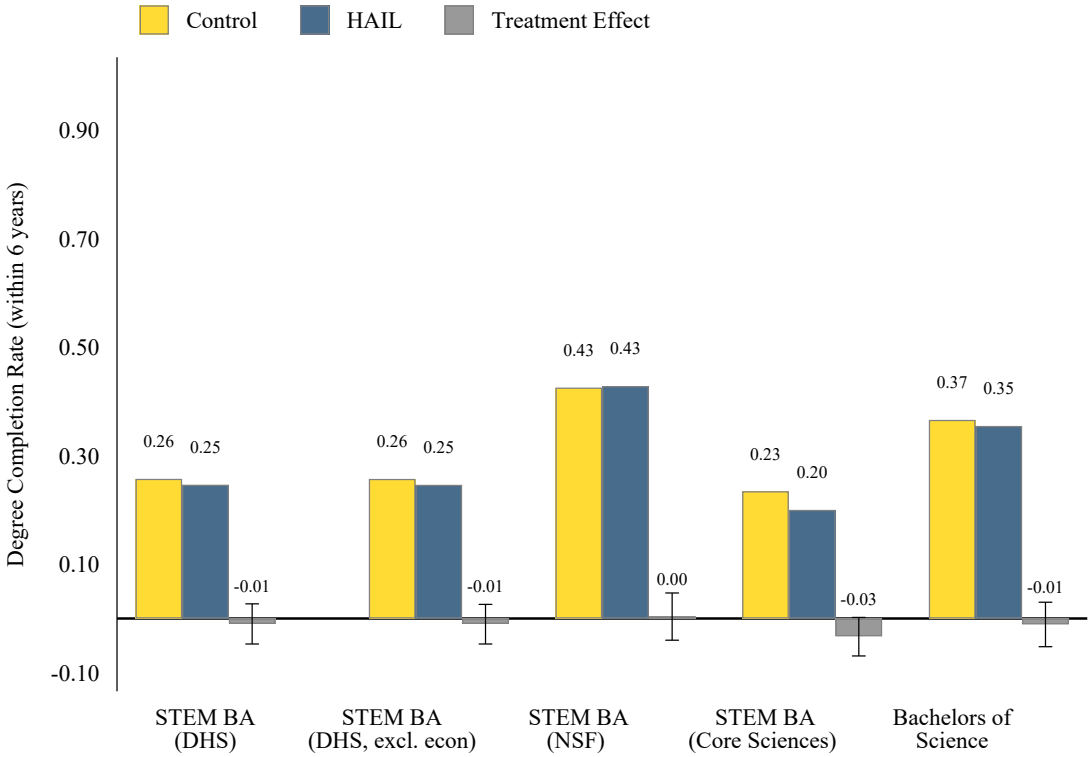
We use degree attainment from NSC, which also we structure into three terms per academic year using dates provided by NSC for degree attainment to align with the enrollment terms. We supplement this with degree attainment from the STARR transcript data. We measure time to degree from expected high school graduation, similar to how we approach enrollment. Degrees "by" a given term include degrees earned in that term, or any prior term. In cases where a student earns multiple bachelor's degrees, we report their first bachelor's degree earned. We measure 4- and 6-year degree attainment in the summer following the 4th and 6th spring. For example, for a student in the first HAIL cohort (expected to graduate high school in 2016), 4-year BA attainment is measured in the summer of 2020 and 6-year BA attainment in the summer of 2022.

B.3 STEM Degrees

Our primary measure of STEM uses the Department of Homeland Security's list of STEM CIP codes (the 2022 STEM Designated Degree Program List, updated as of January 21, 2022). This includes all 2-digit CIP codes classified under the broad categories of biological and biomedical sciences, mathematics and statistics, physical sciences, and engineering, as well as certain 6-digit CIP codes under other categories. We test the robustness of our findings to different definitions of STEM:

- DHS excluding economics: excludes economics from the DHS definition due to the subjectivity of defining only certain economics degrees as "STEM."
- NSF: National Science Foundation's "science and engineering (S&E) fields" (National Science Board, National Science Foundation 2024). Using a crosswalk provided by NCES, we identify all of the CIP codes which crosswalk with these broader categories of STEM, at the 4-digit CIP level, and supplement with "degree major" information from NSC when CIP code is not available.
- Core Sciences: Includes any degree in biological sciences, atmospheric sciences, computer sciences, mathematics, chemistry, physics, and engineering (e.g., Griffith 2010; Riegle-Crumb et al. 2012; Hatfield et al. 2022).
- Bachelor of Science: National Student Clearinghouse degree type is Bachelor of Science.

Appendix Figure 2 shows the results using each of these definitions of STEM. Regardless of which definition we use, we find no effect of HAIL on STEM degree attainment.



Appendix Figure 2
Effect of HAIL on STEM BA degree attainment using alternative STEM definitions

Notes: The first column, "STEM BA (DHS)" is the original definition we use in our current analyses. With the exception of "Bachelor of Science" (which is exclusively from NSC), we start with STEM degree information from the STARR dataset and then merge on NSC data. We count a student as having a STEM degree if they are coded as having a STEM degree in either dataset.

B.4 Projected Earnings

We use data from Opportunity Insights (Chetty et al., 2022), which reports median age-30 income by institution. Chetty et al. (2022) use 2014 IRS tax data for individuals born in 1984, and merge this to college enrollment data. Students are assigned their most-attended institution between ages 18-21. The publicly available data includes median earnings at the institution-level. To mirror the variable measurement, we calculate each student's modal institution attended between year 1 and year 4 following high school, and merge on the institutional earnings from Chetty et al. (2022) using OPEID code to measure *projected earnings* for our sample at age 30. Students with multiple modal institutions are assigned the average earnings for their two modal institutions. We measure median earnings in 2025 values, adjusted for inflation using the consumer price index (CPI-U).

B.5 Institution Quality and Institution Characteristics

To create a continuous measure of institution quality, we calculate the average SAT score of Michigan public high school students who finished high school between 2013-2015 and who enrolled at each institution. We assign HAIL students the average institution quality for the first school they attend in their first and second year following expected high school graduation, measured in the first fall and the second fall respectively, and standardize the measure using the control group.

The remaining institution characteristics are from IPEDS, College Scorecard, and from Opportunity Insights. Median parental income comes from Opportunity Insights (Chetty et al., 2022). Endowment

data comes from the IPEDS finance survey, which we download using Urban Institute's Education Data Portal (U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2016). The remaining characteristics use College Scorecard data (2016). Dollar amounts are measured in 2025 values, adjusted for inflation using the consumer price index (CPI-U). Individuals are assigned an average of the institution character characteristics for their year 1 institution (first institution attended) and their year 2 institution (first institution attended). For students who only enrolled in year 1 following high school, they are assigned the values for their year 1 institution. The same logic is applied for students who only enroll in year 2. Students who do not enroll in years 1 or 2 receive missing values.

B.6 Credits

We construct measures of cumulative credits earned while enrolled at four-year institutions and credits earned while enrolled at community colleges. The transcript data does not allow us to distinguish between transferred credits, AP credits, and dual enrollment credits. Counting these credits would result in counting the same credits multiple times, first at the initial institution and again at the institution students transfer to. Therefore, our measures exclusively focus on credits attempted and earned at a given school.[1] We also focus exclusively on undergraduate credits, using our enrollment measure to exclude any graduate credits. Students not enrolled anywhere in a given term are assigned zero credits. For those who do not attend public colleges within Michigan (i.e. are enrolled, but we do not have transcript data for them), we impute credits attempted and credits earned (described below).

Selection into the Transcript Sample For those students who attend public colleges within Michigan, we have access to student transcript data. While these transcripts are far richer than the NSC data, we only have them for students attending public colleges in Michigan. Students attending private colleges, and those attending college outside Michigan, are missing from these data. This could induce selection bias: the HAIL offer plausibly could have affected the likelihood that a student appears in the STARR data by increasing the share of students attending a public college in Michigan. We therefore use non-missing STARR data combined with universal NSC enrollment data to impute missing credit information for those students who do not attend public colleges within Michigan.

Appendix Table 5 shows the effect of HAIL on selection into the transcript sample by comparing NSC enrollment (with universal coverage) to enrollment calculated from the transcript data. The estimated effect of the HAIL offer on attending a public college in Michigan, and therefore having any credits reported in the transcript data, is 4.1 percentage points, with a standard error of 2.0 percentage points. Since HAIL induced students into the STARR sample, a simple comparison of credits attempted and earned between the treatment and control groups is likely biased in favor of finding a positive treatment effect, since we are more likely to observe credits for students in the treatment group.

Credit Imputation To address this, we impute credits attempted and earned for students who are enrolled in either a private institution within Michigan, or an institution outside the state of Michigan. We follow a similar procedure to that of Angrist, Autor, and Pallais (2022).

For terms where students attend Michigan public institutions, we have their actual credits attempted and earned. For terms where students attend out-of-state institutions or in-state private institutions, we impute credits (attempted and earned) at the student-by-term level, separately for 2- and 4-year enrollments. To do so, we regress credits on enrollment term (fall, spring, summer), year (1-6 post high school graduation), and enrollment status (full-time, half-time, less than half-time, withdrawn/deceased). For four-year enrollments, we also include indicators for Barron's competitive index (highly or most competitive, very competitive or competitive, less-competitive or non-competitive four-year school) and the interaction between Barron's category and enrollment status. We then assign student-terms with missing credits their predicted values from the regression. Credits are imputed for 8.7 percent of student-terms (7.3 percent of student-terms for treated students, 9.2 percent of student-terms for control-group students).

[1] After this cleaning process, 99 percent of student-term observations have 18 or fewer credits attempted or earned, the standard upper-bound for full-time enrollment. However, the cleaning process is imperfect, and one percent of student-terms have credits attempted or earned that exceed 18. We top-code credits in any given semester (regardless of the type of institution) at 22 credits (changing the credits attempted/earned for 0.1 percent of cases). This does not meaningfully change the average credits attempted or earned.

Appendix Table 5
Effect of HAIL on selection into the transcript data sample

Outcome	Effect of HAIL offer (ITT)	Control mean
Enrolled at any college, Year 1 (NSC)	-0.001 (0.014)	0.905
Attempted any credits, Year 1 (transcript data)	0.041 (0.020)	0.730
N school-cohorts	1026	

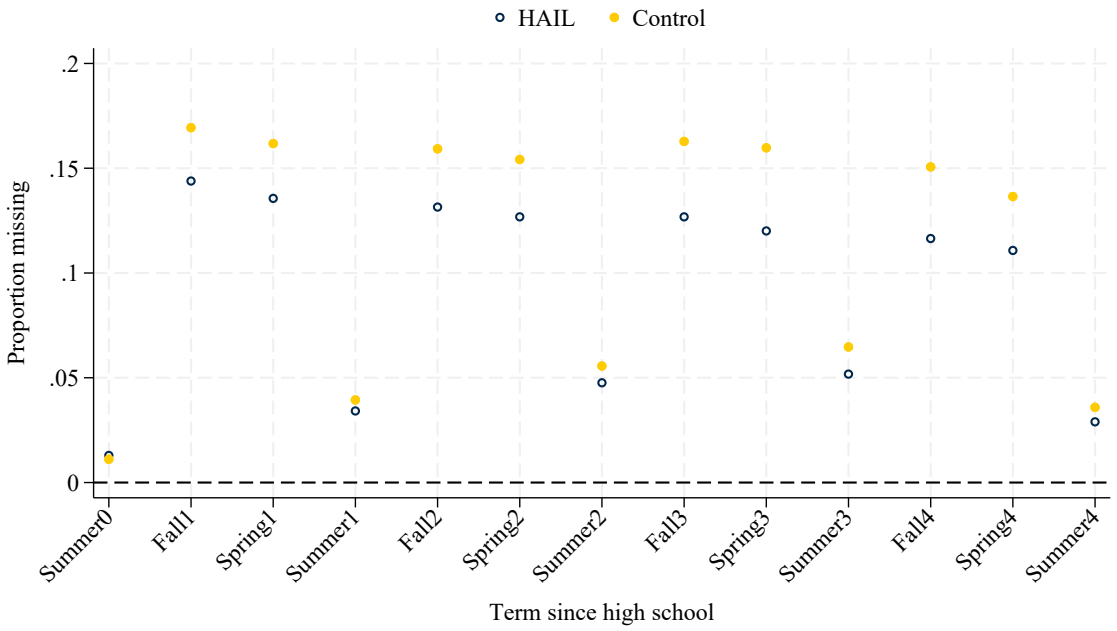
Notes: All analyses done at the high-school-by-cohort-level. Robust standard errors reported in parentheses. Treatment effect coefficients are from estimating Equation (1), the coefficient reports estimates of β_1. Credits calculated at the term-level and include undergraduate coursework taken at a given institution and collapsed to the student-year level. Difference between treatment effect on enrollment in NSC and treatment effect on credits attempted in the transcript data represents selection into the transcript sample.

Lee Bounds In addition to our primary specification, we also estimate Lee Bounds (Lee, 2009). Lee Bounds allow us to bound our treatment effects on credits by modeling the likelihood of missingness from the sample as a latent feature that is predicted by treatment assignment. The lower bound assumes all additional treated students who select into the sample are the top performers, and the upper bound assumes they are the bottom performers. In practice, this means trimming the highest values of earned credits for a portion of the sample equivalent to the differential attrition rate, then estimating treatment effects on the trimmed sample (lower bound); the upper bound is estimated on a sample with the equivalent proportion of the lowest credits trimmed. Lee Bounds do not bound the average treatment effect on all students. Rather, the approach bounds the effect for individuals for whom we would have observed credit outcomes regardless of treatment.

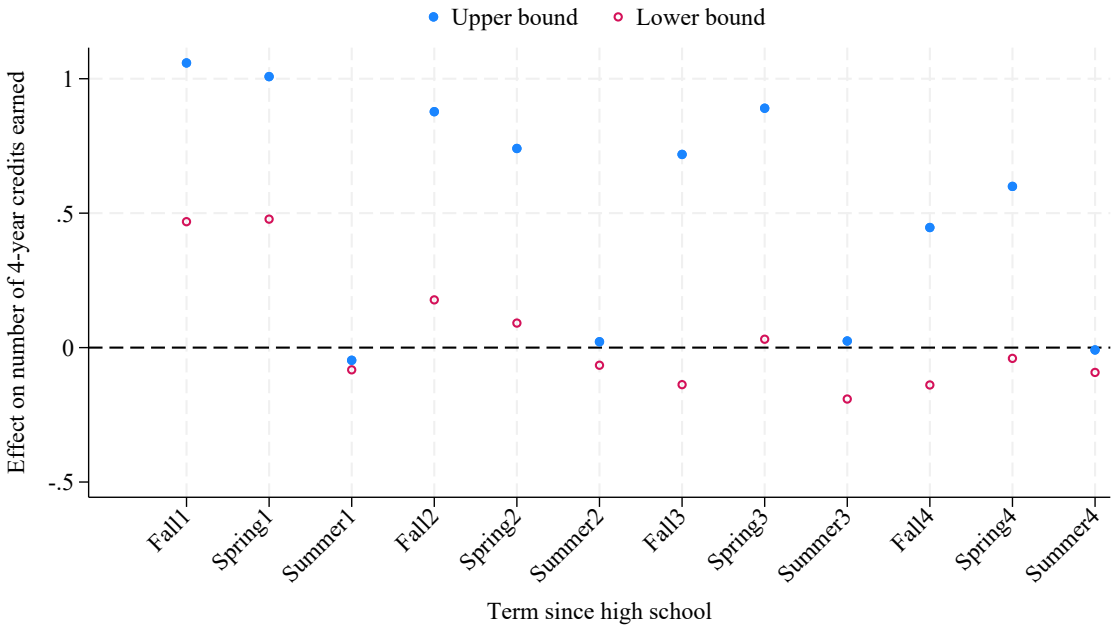
For each semester, we calculate the treatment-control difference in missingness for each credit outcome, at the student-level. We define missing as a term where a student is observed as enrolled in NSC, and missing in the transcript data. In each semester, we drop the top and bottom N_{drop} observations, where N_{drop} is equal to the differential missingness rate times the number of treated students. For example, in Fall 1, the missing rate for control is 16.94 percent, and the rate for treated is 14.39 percent, for a difference of 2.55 percent. There are 1,932 treated students. Thus in Fall 1 N_{drop} = 0.0255 * 1932 = 49 students. Appendix Figure 3 below shows the missingness by treatment status and academic term. The difference between these two is the differential missingness rate.

We estimate treatment effects with the trimmed variables at the school-by-cohort level, using Equation 1 (i.e., controlling for strata and clustering at the high-school-level). The lower bound estimates the effect of HAIL on the credit outcome, dropping the top Ndrop observations from the credit outcome. The upper bound does the same, dropping the lower Ndrop observations.

Appendix Figure 4 shows the bounds on earned credits from four-year institutions in each academic term. The treatment effects using Lee Bounds are largely consistent with the results of our primary approach using imputation: effects on credits earned in fall and spring of years 1 through 4 are likely positive, effects on credits earned in years 5 and 6 are likely negative, and credits effects in summer terms are close to zero.



Appendix Figure 3
Missingness by Treatment Status and Academic Term



Appendix Figure 4
Lee Bounds on Earned Four-Year Credits, by Academic Term: ITTs

B.7 Calculating Control Complier Means

Our LATE estimates tell us the effect of UM attendance (induced by the HAIL intervention) on our outcomes of interest. That is, they assume HAIL's effects on degree attainment are driven by the compliers who were induced by the HAIL offer to attend the UM flagship. To compare these estimates to the relevant counterfactual, we also estimate control complier means (CCMs).

We consider the overall control mean for any outcome of interest to be a weighted average across three types of students: always-takers, never-takers, and compliers. We directly observe outcomes for the always-takers—those students in the control group who enroll at UM. Assuming control and treated always-takers (and never-takers) have the same outcomes in expectation, we also observe the outcomes for the never-takers—those students in the treatment group who do not enroll at UM.

First, we estimate the share of the control group that is in each of the relevant groups. For this to work, we must assume there are no "defiers"—students who are less likely to enroll at UM when they receive a HAIL offer relative to no offer (Imbens and Rubin, 1997). We calculate the proportion of always takers as the share of the control group who enroll at UM:

$$p_{at} = E[EnrollUM|Treat = 0] \tag{1}$$

Similarly, the proportion of never takers is the share of the treatment group who do NOT enroll:

$$p_{nt} = 1 - E[EnrollUM|Treat = 1] \tag{2}$$

Then, the proportion of compliers is simply the remainder (assuming no defiers):

$$p_{cm} = 1 - p_{at} - p_{nt} \tag{3}$$

Next, we can calculate outcome means for always- and never-takers. By assuming their behavior is not affected by treatment, we are assuming control and treated always (and never) takers have the same outcomes in expectation. The always-taker outcome mean is the mean among control students who enroll at UM:

$$\bar{Y}_{at} = E[Y|Treat = 0, EnrollUM = 1] \tag{4}$$

And the never-taker mean is the outcome mean among treated students who do not enroll:

$$\bar{Y}_{nt} = E[Y|Treat = 1, EnrollUM = 0] \tag{5}$$

We can then use the fact that the overall control mean is a weighted average across the three types of students to solve for the CCM as a function of observable quantities:

$$\begin{aligned} \bar{Y}_{control} &= p_{at}\bar{Y}_{at} + p_{nt}\bar{Y}_{nt} + p_{cm}\bar{Y}_{cm} \\ &\Rightarrow \bar{Y}_{cm} = \bar{Y}_{control} - p_{at}\bar{Y}_{at} - p_{nt}\bar{Y}_{nt}p_{cm} \end{aligned} \tag{6}$$

We estimate each of these at the student-level, since compliance is an individual-level behavior. However, we weight all averages by the inverse of the school size to be comparable with our high-school-by-cohort-level treatment effects.

Estimated CCMs for outcomes that should be bounded at zero can be negative, and are in some instances. Treated and control students (and all sub-populations, including compliers, which are defined by pre-treatment characteristics) are by assumption identical in expectation but may not be in our finite sample. Kling, Leibman, and Katz (2007) describe this in detail:

> For binary outcomes, sampling variation can produce negative estimates of the CCM. Our method assumes that there would have been the same fraction of noncompliers in the control group as were observed in the treatment group and that these noncompliers had exactly the same outcome prevalence as treatment noncompliers. In any one sample, this method may produce a negative estimate of the CCM if a particular realization of the treatment noncomplier mean is higher than the realization of the control noncomplier mean, even though the method is unbiased in repeated sampling. (p. 94)