Diversifying the STEM Pipeline:

Evidence from STEM Summer Programs for Underrepresented Youth

Sarah R. Cohodes, Helen Ho, and Silvia C. Robles

Abstract: Pipeline programs intended to increase diversity in STEM fields are common, but there is little rigorous evidence of their efficacy. We fielded a randomized controlled trial to study a suite of such programs targeted to underrepresented high school students and hosted at an elite technical institution. Students offered seats in the STEM summer programs were likelier to enroll in, persist through, and graduate from elite colleges and to graduate with a degree in a STEM field. These improvements in college outcomes raised predicted earnings by 3–15 percent.

Keywords: STEM degree production, inequality, economics of education

JEL Classification: I21, I24, J15, J16

Sarah Cohodes is Associate Professor of Public Policy at the University of Michigan Gerald R. Ford School of Public Policy (scohodes@umich.edu). Helen Ho is a Research and Evaluation Manager at Chinese for Affirmative Action, a nonprofit that advocates for affirmative action and diversity, equity, and inclusion policies. She is an author of this article in her personal, not professional, capacity. Silvia Robles is a Research Scientist at the Annenberg Institute at Brown University. An Online Appendix is included. We are grateful to the staff, applicants and participants of the summer programs that are the subject of this paper and to the host institution and its institutional research office. Because the institution is anonymous, we unfortunately cannot individually thank by name the many people who supported this project. Their commitment was instrumental to this experiment, as was feedback and suggestions from Joshua Angrist, Lawrence Katz, and Amanda Pallais. Rebecca Johnson, Stephanie Owen, Samantha Eyler-Driscoll, and seminar and conference participants at APPAM, AEFP, Colby College, Heinrich Heine University Düsseldorf, NBER, the Opportunity and Inclusive Growth Initiative at the Minneapolis Federal Reserve Bank, the University of Michigan, and Yale University provided helpful comments. Excellent research assistance came from Elizabeth Huffaker, Taylor Myers, Marisa Morin, and Katharine Parham. We thank the Alfred P. Sloan Foundation, the Teachers College Provost's Investment Fund, the National Center for Institutional Diversity at the University of Michigan, the Lab for Economic Applications and Policy at Harvard University, and the National Science Foundation for financial support. The National Bureau of Economic Research, Teachers College Columbia University, Blueprint Labs, Harvard University, and the University of Michigan all provided institutional support. This study is registered in the AEA RCT Registry under identifying number AEARCTR-0002888. It was approved by the Institutional Research Boards at Harvard University, Teachers College Columbia University, the University of Michigan, and the Host Institution. The data and code used in this article are available online: Cohodes, Sarah R. Helen Ho, and Silvia C. Robles. Data and Code for: Diversifying the STEM Pipeline: Evidence from STEM Summer Programs for Underrepresented Youth (ICPSR-246199). Ann Arbor, Michigan. Inter-University Consortium for Political and Social Research. With the permission of the Editor, some variables and code have been omitted to preserve the anonymity of the host institution.

doi:10.3368/jhr.1024-13934R2

I. Introduction

Black and Hispanic workers are underrepresented in the high-wage, college degree–holding science, technology, engineering, and mathematics (STEM) workforce (National Science Board 2021). The lack of diversity in STEM fields contributes to racial and ethnic wage gaps (Budig, Lim, and Hodges 2021), reduces innovation (Parrotta, Pozzoli, and Pytlikova 2014; Hofstra et al. 2020; Yang et al. 2022), and dampens economic growth (Cook, Gerson, and Kuan 2022; Hsieh et al. 2019).i Billions of dollars are spent each year on STEM pipeline programs to correct racial disparities in STEM fields, but we know little about such programs' efficacy.ii

Increasing STEM college degree attainment among underrepresented minorities is necessary to ultimately diversify the STEM workforce. In 2019, approximately 9 percent of STEM bachelor's degrees went to Black students and 16 percent to Hispanic students, despite these groups representing 14 and 21 percent, respectively, of the US college-age population (National Science Board 2022). The disparity in STEM degree attainment is not due to differences in interest. Upon entering college, underrepresented minority students and their white peers plan to major in STEM fields at similar rates, but the former are likelier to switch away from a STEM field or leave college (Riegle-Crumb, King, and Irizarry 2019), which suggests that college experiences are important factors in STEM degree attainment. In particular, access to well-resourced colleges and STEM preparation could help students persist in college and in a STEM major.

Given that access to college and preparation for college STEM experiences are shaped prior to college entrance, STEM-focused enrichment programs for high school students are promising vehicles to reduce disparities in STEM degree attainment and STEM workforce participation. However, the efficacy of such programs in boosting long-term STEM persistence is unknown.

The existing evidence primarily relies on short-term survey assessments and observational studies whose findings may be substantially driven by selection bias (Kitchen, Sonnert, and Sadler 2018a; Kitchen, Sadler, and Sonnert 2018b; Bradford, Beier, and Oswald 2021).[iii] An exception is Robles's (2018) prior investigation of one of the most intensive of the three programs we examine here. Using long-term administrative data on earlier cohorts and a selection-on-observables design, she finds that access to a six-week residential STEM summer program increased matriculation at the host institution, college graduation, and likelihood of graduation with a STEM degree. However, this study does not fully account for selection into the program. Thus, to better understand whether STEM programs for high school students can serve as tools to successfully diversify STEM fields, we conducted a randomized controlled trial (RCT) of a suite of summer programs targeted at boosting the number of underrepresented students in the STEM degree and career pipeline, following students from their application to the programs through college degree attainment.

This study provides the first randomized evidence on the impact of STEM-focused summer programs on college matriculation, completion, and graduation with a STEM degree. It is also one of the few RCTs among *all* the interventions aimed at diversifying the STEM educational pipeline. Over 2,000 high-achieving, STEM-interested high school students across three cohorts were randomized to three STEM-focused programs and a control group in the summer between their junior and senior years of high school in 2014, 2015, and 2016, prior to college application. Given program needs, the randomization design partially took into account program application scores. We thus present both estimates for the subpopulation of completely randomized applicants and estimates subject to additional underlying assumptions for the full sample of conditionally randomized applicants.

The programs were held at the host institution (HI), an elite technical university in the Northeast.[iv] They differed in modality and intensity: The first consisted of six weeks of full-time on-site programming, the second of one week of full-time on-site programming, and the third of six months of periodic online meetings and a short on-site visit. Students were selected into the randomization pool on the basis of their academic preparation and a holistic assessment of need that included whether they had backgrounds that were underrepresented in STEM fields. The six-week program was held on the HI campus and offered a shortened version of the HI's freshman curriculum, along with college counseling, field trips, introductions to role models in STEM fields, and a college-like living experience. The one-week version of the program offered a short, intensive course in a STEM field and an abbreviated version of other aspects of the six-week program, again on the HI campus. Finally, the online version of the program offered a six-month engagement in STEM enrichment activities, with online speakers and interactions and a short "conference" visit to the HI campus over the summer. All programs included some information on the college admissions process, with the six-week program offering personalized counseling, the one-week program providing information sessions, and the online program including a forum where the students could ask college admissions questions.

We use data from the National Student Clearinghouse (NSC) to measure college enrollment and graduation. Almost all students in the control group (88 percent) attended a four-year college immediately after high school graduation, and the programs increased this by a small and statistically nonsignificant magnitude (1–4 percentage points). However, the programs did induce large, statistically significant shifts toward attending elite schools. For the six-week program, these additional elite-institution enrollments were primarily at the HI. For the one-week and online programs, they were split between the HI and other colleges. By the fourth year of study, those offered a seat at a STEM summer program maintained these enrollments.

The programs' shifts in college enrollment translate to higher rates of on-time college graduation. Despite the control group being academically talented, only 56 percent of control students graduated within four years from any four-year school. The STEM programs increased this figure by 3–7 percentage points, though the differences are not statistically significant. The impacts on graduation become more precise at a longer graduation horizon: All three programs boosted the share of students who attained a bachelor's degree (BA) within six years by 2–9 percentage points, though the difference is not statistically significant for the six-week program. BA attainment from elite institutions increased by 9–15 percentage points, with most of the increase for the six-week program coming via the HI and the gains in elite BA attainment from the other two programs being shared across the HI and other institutions.

The degree gains were primarily in STEM fields, particularly engineering, reflecting both an overall increase in the number of degrees earned and a shift to STEM fields among graduates. In the control group, 52 percent of students graduated within six years with a STEM degree—corresponding to 68 percent of the degree recipients in this group. The six-week and one-week programs increased the rate at which students graduated with a STEM degree to 64 percent and the online program to 55 percent (the latter increase is not significant). We summarize the combined effects of these shifts in quality of the institution of enrollment and in choice of major by means of estimates of program impacts on potential earnings, as measured by the College Scorecard for previous cohorts at each college-by-major combination. This exercise implies that being offered a spot in one of the summer STEM programs raised students' potential earnings by 3–15 percent via effects on quality of the degree-granting institution and choice of a STEM major.

At the time of the study, the marginal cost of the six-week program was approximately $15,000 per student, and the one-week and online programs cost approximately $2,000 per student. A back-of-the-envelope cost-effectiveness calculation shows that all three programs have a net positive return due to the predicted increases in student earnings alone—before we even consider other potential benefits to community or student well-being from the programs.

We find evidence that the programs' effect on degree completion is due to the shifts that they induce in the quality of the institutions in which students enrolled. The increases in overall graduation rates are on average similar to what would be predicted by these shifts alone, as measured by institution-level graduation rates. The programs potentially achieved these shifts by enhancing students' information about colleges and the college application process, as shown via administrative application data from the HI and survey data on other college applications. We also use the survey data to explore other mechanisms, such as improvements in study and independent living skills and high school preparation. The programs did boost Advanced Placement/International Baccalaureate (AP/IB) enrollment in high schools and study and life skills, but the magnitudes of these changes are not large. Thus, we suggest that college quality likely plays the greatest role in explaining the graduation effects that we identify. We find some potential evidence for the role of signaling in college admissions for the six-week program at the HI, but we conclude from the full body of evidence on mechanisms that the college application channel is the most important.

This paper makes three main contributions. First, we add to the evidence on STEM degree attainment and diversity among STEM degree holders. Most research on STEM degree production focuses on what happens during college, concentrating on the gender or race match between students and instructors or peers (see, for example, Bettinger and Long 2005; Hoffmann and Oreopoulos 2009; Griffith 2010; Carrell, Page, and West 2010; Bettinger 2010; Price 2010; Fairlie, Hoffmann, and Oreopoulos 2014; Fischer 2017; Griffith and Main 2019), student beliefs in their own capability and signals from grades (Astorne-Figari and Speer 2019; Kaganovich, Taylor, and Xiao 2021; Owen 2023; Kugler, Tinsley, and Ukhaneva 2021; Owen 2021), and institutional effects (Griffith 2010; Arcidiacono, Aucejo, and Hotz 2016). Less attention has been paid to the preparatory experiences that may shape college attendance and major choices and eventually occupational sorting and labor market disparities, despite the potential influence of precollege experiences on STEM degree attainment (Sass 2015; Green and Sanderson 2018). The few studies on high school STEM exposure find differing effects on STEM major choices and degree attainment. For the US, Darolia et al. (2020) find that exposure to more STEM courses in high school does not increase STEM degree attainment in college, but Nomi, DeChane, and Podgursky (2025) find that access to Project Lead the Way (a STEM program) increased STEM major declarations (the STEM major outcome available in their timeframe). Outside of the US, De Philippis (2021) finds that, in the United Kingdom, high school STEM exposure increases the likelihood of male students majoring in STEM, and Joensen and Nielsen (2016) find an increase only for female students in Denmark. Although the differences in these findings may be due to differences in context, it is also possible that broad programs with no specific focus on underrepresented students or those that do not affect college applications may have little effect.

Second, we contribute to the understanding of access to college, the match between student preparation and quality of institution of enrollment, and the potential for college education to reduce racial and economic inequality in the US. There are large gaps in college enrollment by family income in the US (Bailey and Dynarski 2011; Chetty et al. 2020; Dynarski et al. 2021). In addition, there are differences in the type and quality of the institutions where students of different socioeconomic groups enroll (Gerber and Cheung 2008; Baker, Klasik, and Reardon 2018). Even high-achieving underserved students lag in their college enrollment and selectivity (Hoxby and Avery 2013; Dillon and Smith 2017), which results in "undermatch," whereby students who could succeed at selective institutions do not even apply and thus cannot enroll. The college a person attends can influence her likelihood of (on-time) graduation and of graduation with certain degrees, along with her future employment and earnings (Hoekstra 2009; Cohodes and Goodman 2014; Zimmerman 2014; Goodman, Hurwitz, and Smith 2017; Chetty et al. 2020; Bleemer 2021). Interventions prior to college application can influence enrollment and the specific institutions where students enroll (Avery 2010; Avery 2013; Carrell and Sacerdote 2017; Castleman and Goodman 2018; Andrews, Imberman, and Lovenheim 2020; Dynarski et al. 2021; Barr and Castleman 2025). Similarly to effective college counseling and informational interventions that modify college application and enrollment behavior, the STEM summer programs we examine happened at a crucial time: when the students were seriously considering college but had not yet applied. However, the STEM summer programs we focus on differ from many college access programs in their intensity and focus on STEM.

Finally, this paper is relevant to a large literature on the impacts (or lack thereof) of affirmative action in college admissions. STEM summer programs do not introduce group-based preferences in college admissions—the policies typically in focus in the affirmative action literature in economics. They do, however, focus on populations characterized by a broad definition of need that includes belonging to historically underrepresented groups, and they aim to increase access to STEM fields and elite universities for these groups. Much of the literature on affirmative action is concerned with mismatch—the idea that underrepresented minority (URM) students will be unprepared for the academic rigor of campuses with affirmative action preferences and thus might be made worse off by such policies (see Arcidiacono and Lovenheim 2016 for an overview). Although Arcidiacono, Aucejo, and Hotz (2016) find some evidence of mismatch, Bleemer (2022) finds college and earnings benefits for URM students induced to attend more selective University of California campuses because of affirmative action. Our work adds to the evidence that when URM students are induced to attend high-quality institutions, they successfully reap the benefits of those institutions, in contrast to the predictions of mismatch theory.

Overall, this paper shows benefits from enrollment in STEM summer programs in terms of graduation from elite colleges and in STEM fields, highlighting the potential for specialized programming to improve outcomes for underrepresented students and increase diversity in the STEM pipeline. The benefits appear to operate through an increase college quality. For the six-week program, the gains in college quality primarily come from the HI, implying that the benefits of the flagship program are reaped at the HI. However, both the one-week and online programs improve college quality through enrollments at the HI *and* other elite campuses. This means that the programming created and operated by the HI improves outcomes both by shifting enrollment to the HI and by increasing enrollment at other elite institutions. This implies that such programming is likely underprovided, as its benefits accrue beyond the provider. Timing interventions so they have the potential to change the higher education institution attended is a promising path for increasing college quality and graduation.

The paper proceeds as follows. Section II describes the program background and context, including more details on the interventions. Section III details the data, and Section IV explains the study design and estimation methods. Our results are reported in Section V, with a discussion of potential mechanisms in Section VI. Section VII concludes.

II. STEM Summer Programs at the HI

The HI maintains an office devoted to outreach programs to increase representation of URM students in STEM fields; we refer to this unit as the outreach office. Its programming includes outreach to the local community with initiatives designed for elementary and secondary students and national summer programs for high school juniors. The summer programs are the focus of this study. The aim of the programs is to diversify the STEM workforce and increase access to STEM careers by exposing students to high-achieving peers, STEM mentors, a STEM curriculum, tours of a college campus and research facilities, and college admissions information. Instructional staff in the programs are alumni, lecturers at the HI, and high school teachers but not typically tenure-line faculty at the HI. Recruitment is national.

All programs cover student costs except for transportation to and from the HI. This includes (implicit) tuition, dormitory lodgings, and meals. The programs are funded by the HI, with some funding from earmarked charitable gifts. High-achieving students in any geographic region can be recruited as long as they are US citizens or permanent residents. One source of student information used in direct mailings for recruitment is the PSAT; students need not submit test scores to apply to the programs, but most do, and the scores are used in applicant rankings. Most students who attend the summer program are URM, but URM status is not a requirement of the program, as we detail in Section IV.A.

We describe each summer program below as it existed in the summers of 2014–2016, the period over which the randomization occurred. All of the outreach office's programs offer similar experiences designed to promote persistence in STEM fields, but the intensity and modality of the experiences vary.

1. *Six-week program*: The six-week program is the longest-running summer program of the three studied. It is a residential program that immerses rising high school seniors in rigorous science and engineering classes. Students take courses in math, physics, life sciences, and humanities, as well as a STEM-related elective course with topics ranging from digital design to genomics. In addition, students take tours of labs and workspaces at the HI, attend workshops with industry leaders, academics, and admissions officers, and interact with teaching assistants who are current college students. Students also visit STEM-focused companies and workplaces. The program encourages social cohesion by bringing students together to live in dorms at the HI and leading team-building exercises. Approximately 80 students are offered a seat in this program each year.
2. *One-week program*: The one-week program encapsulated some aspects of the six-week program in a shorter time frame and was also a residential program. Over the week, students completed a project course in an engineering field, attended admissions and financial aid sessions, toured labs, met with HI faculty, students, and alumni, and participated in social events. The time constraint necessarily reduced the dosage of all aspects of the six-week program, though to what extent outcomes are sensitive to this reduction is an empirical question. Typically, 75–120 students participated in this program each year.[v]
3. *Online program*: The online treatment draws on communications technology to serve students. This six-month program provides a platform for multimedia interaction between students and instructors, HI staff, and industry leaders. HI students are hired to mentor small groups of participants and lead discussions. The online summer program provides top-down content in the form of videos, articles, or webinars. Students must also complete project-based engineering assignments. The forum and discussion groups provide user-generated

(and instructor-facilitated) content. The online program does not include formal college counseling, unlike the six- and one-week programs, but includes a discussion forum dedicated to the college admissions process with participation from HI admissions and financial aid officers to answer questions. Students also spend five days on campus presenting their final projects, attending workshops, and meeting their classmates in person. The campus visit occurs five weeks into the online experience, which lasts until the end of the calendar year. Approximately 150–175 students participate in the online program each year. The summer sessions that we study occurred well before the COVID-19 pandemic, but the technology platform used for the online program facilitated a transition to digital learning for all of the summer programs in COVID-affected years.

4. *Control condition*: Students assigned to the control condition also applied to the HI's summer programs but were not randomly assigned to participate in any programs offered by the outreach office. However, these applicants were generally also accomplished students (typically in the top third of applicants to the summer programs as a whole). Being assigned to the control group does not mean that the applicant had no exposure to STEM-focused programming. Many students in this group participated in alternative summer programming, both STEM-focused and otherwise, such as programs administered by other universities or organizations such as Girls Who Code or Leadership Enterprise for a Diverse America. Specifically, among the 65 percent of the control group students who responded to our survey, 46 percent participated in summer programming (including STEM), with 29 percent of the respondent control group participating in STEM-focused summer programs. The remaining respondent control group students worked or studied over the summer in lieu of enrolling in specialized summer programs.

III. Data and Descriptive Statistics

This study measures the effect of STEM summer programs on the college outcomes of high-achieving underrepresented high school students using data from an experiment that randomized admission to these programs in the summers of 2014, 2015, and 2016. Below, we detail the data sources used, which include records from the outreach office on summer program application and admission, college attendance and graduation records, and survey data, and describe the characteristics of program applicants.

A. Data

The data for our main analyses come from two main sources: program application and admissions information from the HI and college attendance and graduation information from the NSC. All applicants in the three cohorts from 2014 to 2016 were admitted via conditional random assignment; the assignment and randomization process was jointly created by the research team, program staff, and the HI institutional research office to meet both research and operational needs (Cohodes, Ho, and Robles 2022). Background information on the randomized sample comes from program applications, which included demographic information, academic qualifications, essays, and a baseline survey. The outreach office provided details on the admissions process, the ratings of applicants' files, offers of admission, and which students ultimately participated in the programs.

The outcome data come from records of college enrollment provided by the HI institutional research office and the NSC; the former also provided information on applications and majors at the HI. Almost all individuals in the applicant pool appear in the HI or NSC college enrollment data, and because all students' information was shared with the NSC and HI for matching to enrollment data, there is no differential attrition in the possibility of appearing in the college data (see Online Appendix Table A.3). Our time horizon allows us to observe graduation within six years for all three cohorts. The NSC data also include information on students' majors, which we categorize as either STEM or non-STEM.[vi] We also categorize college institutions as elite if they are ranked "most competitive" by *Barron's Guide*. Online Appendix Figure A.1 shows the timing of the observations through college for each of the three cohorts, assuming on-time progression.

B. Surveys

Survey data were also periodically collected from the study sample. Longer surveys were conducted (1) in the fall shortly after the summer program, (2) in May of the students' senior year of high school, and (3) in the spring of the students' sophomore year of college. Shorter, more frequent surveys kept track of college enrollment and students' ultimate or intended college major. Respondents received Amazon gift cards if they participated (not contingent on their answering all questions or on treatment assignment), with larger incentives for participation in the longer surveys ($25) and smaller incentives for the shorter ones ($10).

The survey at the end of the summer program included questions about college plans, knowledge of the application process, intended major, and study and life skills. The survey in May of senior year collected information on college application and admission and fall plans. The final long-format survey at the end of sophomore year in college asked about college experiences, majors, and career intentions. More details on these surveys are in Online Appendix C. When a survey included multiple items on similar topics, we constructed standardized indices of the outcome measures from the surveys by outcome "family" using the method in Anderson (2008) to minimize concerns about multiple hypothesis testing.

Response rates were relatively high but declined over time and were lower for the control group. Online Appendix Table A.3 shows the response rates for each survey. For example, for the first long follow-up survey in the fall after program participation, the treatment groups' response rates ranged from 85 to 90 percent; 65 percent of the control group responded to this survey. These differences in response rates are unsurprising given that treatment students were likelier to have positive associations with the program office or be enrolled in the HI, the entity sending out the surveys.

To assess the representativeness of the survey response sample, the analysis compares program impacts on college attendance and graduation between the survey sample and the full sample both with and without inverse propensity-to-respond weights (see Online Appendix Table E.2). Restricting to the survey sample and adding inverse probability weights have an inconsistent impact on the treatment effect estimates. In some cases, these changes to the specification bring the estimates of the impacts on college outcomes for survey responders closer to those for the full sample than to those for the unweighted sample of respondents; in others, they make the estimates for survey respondents look more divergent than those for the full sample. Thus, it is not clear that propensity to respond predicts the program effects in a meaningful way. Following Dutz et al. (2021), we use the unweighted survey responses but caution that the sample is not fully representative. We thus consider our findings based on the surveys to be suggestive.

C. Descriptive Statistics

Table 1 reports demographic and background academic information for the randomized sample. As described below, the randomization included design strata with different probabilities of program offers, so we do not expect the treatment and control groups to be exactly similar on all characteristics. We show in Columns 6–8 that after strata adjustment, the treatment group has few differences from the control group.[vii] Overall, almost all randomized applicants identified themselves as belonging to a group underrepresented in higher education and STEM fields, with

35 percent of the sample being Black, 43 percent Hispanic, and 4 percent Native American (note that these categories overlap, as the students could report more than one race or ethnicity). Approximately one-quarter of the group were first-generation college students, which we define as a student's having no parent who ever attended a four-year college. The program applicants had strong academic backgrounds. The average grade point average (GPA) is 3.86 on a four-point scale, and the average standardized math test score is two standard deviations above the national mean.[viii] The largest contrast between the treatment and control groups is on gender. The outreach office seeks to host programs equally split between young women and men, but the applicant pool skews male. Thus, the randomization strata included gender.

IV. Research Design

The STEM summer program applicants were randomized into receiving an offer to participate in one of the three summer programs or being assigned to a control group. To meet programmatic needs, the random assignment was partially conditioned on applicant rating. This section describes the process through which applicants were randomized and how we estimate the program effects on the basis of that randomization process. Online Appendix A provides further details on the selection and randomization process.

A. Selection

Selection into the programs was a multistep process. Each year, after an initial screening, the outreach office sent approximately 600–750 highly qualified applicants to a selection committee made up of stakeholders, community members, and affiliates with long-standing ties to the outreach office. The selection committee ranked applicants in terms of suitability for the six-week program and provided detailed scores on their academic preparation and personal circumstances considering their grades, test scores, letters of recommendation, and application essays. In addition, because the mission of the programs is to promote access to STEM for traditionally underrepresented populations, the selection process included consideration of the following factors on a holistic basis, though no element in isolation guaranteed admission:

1. The applicant would be the first in the family to attend college
2. The applicant's family did not have science or engineering backgrounds
3. The applicant's high school had historically sent less than 50 percent of its graduates to four-year colleges
4. The applicant attended a high school whose students might face challenges in succeeding at an elite, urban university (for instance, a rural high school or a high school with a predominantly URM population)
5. The applicant was a member of a group underrepresented in the study and fields of science and engineering (African American or Black, Hispanic or Latino/a, or Native American)

The selection committees were idiosyncratic in how they evaluated applications; for example, some committees emphasized academic preparation, and others were more focused on perceived need for the program.

The outreach office also requested regional priorities to increase representation across the country, and these entered into the rankings in 2015 and 2016. In 2014, the outreach office exempted several applicants from randomization and offered them admission to support national representation; we call these cases "certainty spots" and exclude them from the analytic sample. The HI institutional research office performed the final program assignment by lottery after all students were ranked, creating a ranking variable that was a weighted average of the applicant ratings and regional priorities; the rankings were used to allocate students to random assignment blocks as described in Section IV.B below. Notably, because the randomization occurred in a group designated as top applicants from a pool of more than 2,000 applicants, the program effects estimated here may not apply to all applicants or to others outside the selective applicant pool. The program, if offered to rising high school seniors in general, might have very different (or no) impacts on those less academically prepared.

B. Randomization

The selection process above created a pool of 600–750 applicants eligible for randomization each year. Because the outreach office wanted to ensure that top-ranked students had access to one of the summer programs, the study employed a block randomization design. The HI institutional research office placed students into randomization blocks according to a rating variable that took into account application ratings and regional priorities. An overview of the randomization scheme appears in Figure 1 and is described in detail in Online Appendix A.

In general, the highest-ranked students were placed in Block 1 and randomized between the three summer programs. To maintain gender balance in the program, there were different rating score cutoffs for male and female students. Additionally, to ensure students in Block 1 were prepared to take on the rigorous coursework in the six-week program, a math test score floor was imposed for assignment to Block 1.[ix] The remaining students were placed in Block 2 and were randomly assigned between the online program and a control group. The size of the blocks and assignments to programs varied by year on the basis of operational considerations. This randomization scheme formed the research design in cohorts 2 and 3 of the experiment. Cohort 1, in 2014, was subject to a slightly different design, in which students were differentiated to a greater extent and randomly assigned within three blocks. Our results are very similar whether we include or exclude the first cohort (see Online Appendix Table C.2). We include the 2014 cohort to increase statistical precision, despite the slightly different underlying randomization structure, which we account for via inclusion of randomization strata.

The crucial component of the conditional randomization design is the overlap of the online program across the blocks, which we use to extrapolate comparisons between Block 1 programs and the control group in Block 2. The key assumption behind this extrapolation is that if we control for randomization strata (based on application year, gender, block, and regional preferences) and have constant treatment effects, we fully account for differences across the blocks and can compare the applicants assigned to Block 1 (which has no control group) to those in Block 2 (which has a control group). By design, we can fully account for membership in a block with known variables. To test this assumption, we compare our estimates for the conditionally randomized sample to those for the embedded purely randomized sample and find generally similar results.

C. Estimation

We use random assignment to program offers to estimate in two different ways the causal effect of assignment to one of the three STEM summer programs. First, we compare *within* randomization blocks, separately comparing the outcomes for those in the residential programs to those assigned to the online program and the outcomes for students in the online program to those for the control group, as in Cohodes et al. (2024). This comparison relies only on random assignment, with no extrapolation across blocks. However, these comparisons have reduced power given the smaller sample sizes. Second, we follow the structure of the conditional random assignment, comparing the outcomes for students across all of the programs to those of the control group, controlling for blocks. This strategy allows us to use all of the data and increase statistical power but relies on our assuming that treatment effects are constant and that we can extrapolate across blocks.

Using random assignment, we make two comparisons. The first is within Block 1. We estimate the causal effect of assignment to a residential STEM summer program relative to the outcomes of students assigned to the online program:

$$Y_i = \beta_{6w} 6week_i + \beta_{1w} 1week_i + \sum_{j=1}^{J} \theta_j S_{ij} + X_i' \gamma + \epsilon_i. \quad (1)$$

Next, we make use of the random assignment within Block 2, estimating the causal effect of assignment to the online program relative to the control group's outcomes:

$$Y_i = \beta_{on} Online_i + \sum_{j=1}^{J} \lambda_j S_{ij} + X_i' \omega + \eta_i. \quad (2)$$

In both equations, we are interested in a college outcome Y_i such as college graduation for applicant i. Random treatment assignment is shown with the indicator variables $6week_i$ and $1week_i$ in Equation 1 and $Online_i$ in Equation 2. Intent-to-treat estimates of the effects of program offers are shown by the β coefficients. We use intent-to-treat estimates because most students attended their assigned program if offered a spot, with 87 percent of students accepting a seat in the six-week program, 85 percent in the one-week program, and 77 percent in the online program (Online Appendix Table A.2). Very few students ended up participating in a different program, and the outreach office did not offer any spots to students in the control group. β_{6w} and β_{1w} are the estimates of the difference between applicants offered a seat in one of the residential programs and those offered a slot in the primarily online program. β_{on} is the estimate for the difference between applicants offered a spot in the online program and those assigned to a control group. A vector of student-level control variables X_i, including GPA, standardized math scores, race/ethnicity, and free or reduced-price lunch status, increases precision. We include randomization strata, S_{ij}, to account for differences across program years, the preference for gender-balanced enrollment, and the regional preferences.[x] The program offers were randomized within these strata. We use heteroskedasticity-robust standard errors. We refer to the estimates derived with this strategy as the "pure randomization" estimates.

To compare the estimates from Block 1 directly to those for the control group, we use seemingly unrelated regression (SUR) to combine the Block 1 and Block 2 estimates. Additionally, we present estimates of the "Pooled" treatment effect calculated using the Block 1 and Block 2 estimates via SUR. This estimate is the weighted average of the three program effects vs. the outcomes in the control group, with each program estimate weighted by share of the treatment group.[xi] The interpretation of these estimates as the true program effects relies on the same assumptions that we detail below for our second estimation strategy.

Our second estimation strategy is to jointly estimate the treatment effects in comparison to the control group outcomes but relies on additional assumptions. We estimate the impacts from conditional random assignment as follows:

$$Y_i = \tau_{6w} 6week_i + \tau_{1w} 1week_i + \tau_{on} Online_i + \sum_{k=1}^{K} \psi_k R_{ik} + X_i' \phi + \mu_i, \quad (3)$$

where Y_i is an outcome of interest for applicant i such as enrollment in a top-ranked college and $6week_i$, $1week_i$, and $Online_i$ are the indicators for random assignment to an offer of treatment in each of the three programs. The parameters of interest are the τ coefficients, which reflect the intent-to-treat estimate of the effect of assignment to one of the three programs. As above, we include a vector of student characteristics X_i to increase precision and use heteroskedasticity-robust standard errors.

Key to our estimation strategy is the inclusion of a set of control variables, or risk sets, R_{ik}, which are indicators representing the randomization strata. The randomization strata comprise gender, regional preferences, and randomization block and are formed within randomization year.[xii] Offers were randomized within these strata. Students were assigned to randomization blocks on the basis of a rating variable and a standardized test score floor. The rating variable includes ratings by a selection committee, assessment by the HI admissions office, and prioritization for certain regions and states of the country (typically to ensure that the participants were broadly representative of the US as a whole). The conditional randomization estimation method compares students within the same cohort, gender, regional preference, and randomization block.

The fundamental assumption behind our second estimation strategy is that by conditioning on randomization block, we control for all differences across blocks and can compare students randomly assigned to a treatment group to those in the control group even when we do not have a direct treatment–control contrast in the same block. This assumption is feasible if we also assume constant treatment effects, as we discuss further below. Including the online program in both randomization blocks provides the link that makes it possible to draw this comparison. Because block assignment was based completely on known, observable variables, controlling for randomization block should control for all differences between the two groups, assuming constant treatment effects. We thus refer to our estimates derived with this strategy as the "conditional randomization" estimates.

Some potential outcomes notation makes clear the role of constant treatment effects. Each individual has two potential outcomes: the outcome under treatment, Y_{1i}, where 1 indicates treatment assignment, and Y_{0i}, the outcome when treatment assignment did not occur. Treatment, indicated by D, occurs if D_{6w}=1, D_{1w}=1, or D_{on}=1, and does not occur if D=0:

$$Y_{di} = \begin{cases} 1 \text{ if } D_{6w} = 1 \text{ or } D_{1w} = 1 \text{ or } D_{on} = 1 \\ 0 \text{ if } D = 0. \end{cases}$$

There is then an unobservable individual treatment effect: $\delta_i = Y_{1i} - Y_{0i}$. Once we condition on randomization strata, R_{ik}, we have pure randomization within block.[xiii] Because we have pure randomization between treatment and control for the online program vs. the control group in Block 2, and for the six-week and one-week programs vs. the online program in Block 1, by the conditional independence assumption:

$$Y_i \perp\!\!\!\perp D_{on_i} \mid \text{Block } 2, R_{ik},$$

$$Y_i \perp\!\!\!\perp D_{6w_i,1w_i} \mid \text{Block } 1, R_{ik}.$$

We can then write the average treatment effect[xiv] (ATE) for the online program vs. the control group in Block 2 as:

$$\begin{aligned} ATE_{on|Block\ 2,R_k} &= \delta_{on|Block\ 2,R_k} \\ &= \mathrm{E}[\mathrm{Y}_1 \mid \mathrm{D}_{on} = 1, \text{Block 2}, \mathrm{R}_k] - \mathrm{E}[\mathrm{Y}_0 \mid \mathrm{D}_{on} = 0, \text{Block 2}, \mathrm{R}_k] \\ &= E[Y_{\text{on}} \mid Block\ 2, R_k] - E[Y_0 \mid Block\ 2, R_k]. \end{aligned}$$

Similarly to the above, the treatment effect of the six-week program vs. the online program can be written as:

$$\begin{aligned} ATE_{6w|Block\ 1,R_k} &= \delta_{6w|Block\ 1,R_k} \\ &= E[Y_1 \mid D_{6w} = 1, Block\ 1, R_k] - E[Y_{\text{on}} \mid D_{on} = 1, Block\ 1, R_k] \\ &= E[Y_{\text{6w}} \mid Block\ 1, R_k] - E[Y_{\text{on}} \mid Block\ 1, R_k]. \end{aligned}$$

Returning the online program effect, we would like to extrapolate from Block 2 to Block 1 and can do so when the ATEs for the online program are constant across the blocks. Then,

$$\delta_{on_i} = \delta_{on}\ \forall_i \qquad \Longrightarrow \qquad \delta_{on} = E[Y_{on} - Y_0 \mid Block\ 1, R_k] = E[Y_{on} - Y_0 \mid Block\ 2, R_k].$$

We also assume that the ATE of the six-week program has a specific functional form—that is, that it is additive and separable from the effect of the online program. If this assumption holds, we can build a treatment effect for the six-week program vs. the control group:

$$\mathrm{AT}E_{6w} = \delta_{6w|Block\ 1,\, R_k} + \delta_{on|Block\ 1,\, R_k}$$

$$= \delta_{6w|Block\ 1, R_k} + \delta_{on|Block\ 2, R_k}$$

$$= E[Y_{6w} \mid Block\ 1, R_k] - E[Y_{on} \mid Block\ 1, R_k]$$

$$+E[Y_{on} \mid Block2, R_k] - E[Y_0 \mid Block2, R_k]$$

$$= E[Y_{6w} \mid Block\ 1, R_k] - E[Y_0 \mid Block\ 2, R_k].$$

The cancellation of the online treatment effects in the final line of the above holds under constant treatment effects.

Treatment effects for the online program that varied across blocks would threaten this estimation method. In Section V.E.3, we consider in detail what would happen if the constant treatment effect assumption failed to hold. In this case, our estimates would be biased because we could not make the extrapolation across blocks. However, we find limited evidence of bias.

This potential outcomes framework also makes clear the importance of the Block 2 control group presenting a good comparison for Block 1 (after adjustment for overall differences between the two via randomization strata, such as differences in the proportion of each block that is female). Because we cannot observe the behavior of a control group that does not exist, we cannot explicitly test how good a job the Block 2 control group does at simulating Block 1 control behavior. Specifically, if hypothetical Block 1 control group members would have performed worse than the performance difference accounted for by the randomization strata, we would not fully capture that treatment effect because of the lack of higher-rated students in the control group, which would bias the program effect downward. This might be the case if higher-rated control group students would have been less successful in counterfactual schools than the performance difference captured by the randomization strata, perhaps due to lack of services and outreach. Alternatively, if hypothetical Block 1 control group students would have performed better in counterfactual institutions (to an extent not captured by the randomization strata) due to better match, our estimates would be biased upward because some of what we would be attributing to the program estimate would be due to lack of a comparable control group. Because these hypotheticals involve a group that does not exist in our data, we cannot test them.

Section V.E presents several additional robustness checks. We consider multiple alternative specifications and conduct an exercise in which we omit each cohort in turn. Because the randomization possibilities are constrained to only those students within the strata discussed above (that is, complete randomization is not possible), we also present our estimates derived by means of randomization inference to account for the possibility that only a subset of potential outcomes are possible given the constraints of our design (see Athey and Imbens 2017 for a discussion of this approach). Specifically, we rerun our randomization scheme 1,000 times, subject to the same rules and constraints as in the actual randomization. We then compare our point estimate to the distribution of estimates generated by random assignment. As in Fisher's exact test, if the point estimate exceeds the 97.5th percentile of the distribution of randomized estimates (or 95th percentile with a one-tailed test), we consider that estimate to be different from zero. All of these exercises yield results similar to our main estimates.

V. Results

This section details how the programs affected college enrollment, college graduation, and college quality. All of our outcomes are based on windows from the time of high school graduation. For example, six-year college graduation reflects college completion within six years of high school graduation (seven years after the summer program), not six years from the time of college entrance. The tables report the estimates from both the pure randomization (Panels A and B of the relevant tables), SUR (Panel C), and conditional randomization (Panel D) strategies. For conciseness, the figures report only the conditional randomization estimates.

A. College Enrollment and Persistence

We present our key findings on the college trajectory in Table 2, which includes both the pure and conditional randomization estimates for initial college attendance and BA receipt within six years for all four-year colleges, elite colleges, and the HI. Figure 2 reports the attendance and graduation effects of the conditional random assignment, divided among the HI (light shading), other elite institutions (medium shading), and other four-year institutions (dark shading). Complete estimates for all years of college attendance and graduation, at colleges of multiple types, appear in Online Appendix Tables B.1–B.7.

Almost all students in the study sample attended four-year college immediately after high school graduation (one academic year after the summer program), with 88 percent of the control group enrolling in the first year (Column 1 of Table 2). The remaining students enrolled in two-year institutions, joined the military, or worked. Attendance of one of the STEM summer programs has positive, but not statistically significant, impacts on four-year college attendance.

Overall enrollment changed little, but the programs increased enrollment in elite colleges. Unsurprisingly, the programs induced students to attend the HI (Column 2). Approximately 6 percent of the control group enrolled in the HI; an offer of online program admission increased this figure by approximately four percentage points, according to both the pure (Panel B) and conditional (Panel D) randomization estimates. An offer of admission to the six-week program increased HI enrollment by 11 percentage points above the share in the online comparison group and by 15–17 percentage points over the share in the control comparison. This means over 20 percent of students with six-week program offers enrolled in the HI. The one-week program also increased HI enrollment by 4 percentage points over the level in the online comparison group and by 6–8 percentage points over the level in the control comparison, though these results are not consistently statistically significant. These results also serve to illustrate that the pure and conditional randomization estimates are very similar. To summarize, the pooled estimate generated via SUR shows that assignment to HI STEM summer programs increased attendance of the HI by 8 percentage points.

The shifts to the HI are part of an overall shift toward enrollment in selective colleges. This can also be seen in Column 3, which reports enrollment at elite institutions, including the HI. We define elite institutions as those that *Barron's Guide* rates "most competitive," its highest rating. An admission offer to the online program increased enrollment in such institutions by 9 percentage points, which means that enrollment increased at elite institutions *beyond* the HI (as the HI estimate is 4 percentage points). The six- and one-week programs boosted elite college enrollment by an additional 4–7 percentage points over the effect of the online program, resulting in enrollment rates 13–16 percentage points higher than those in the control group. As can be seen in Figure 2, for the six-week program, this shift to enrollment in higher-quality institutions is driven by enrollments at the HI. However, the one-week and online programs increased enrollment at both the HI *and* other elite institutions. The pooled treatment effect is a 12.5-percentage-point increase in elite enrollment, which includes enrollment beyond the HI (as the pooled coefficient for the HI is 8 percentage points). The effects on attendance throughout college are generally similar to the effects on initial enrollment (Online Appendix Tables B.2–B.4).

By their fourth year of college, the STEM summer program participants showed an even greater edge. Students assigned to a summer program maintained their enrollment in the fourth year, whereas control students were less likely to have remained enrolled, though this difference is statistically significant only for the one-week program and pooled estimate (Online Appendix Table B.4). The differences in enrollment reflect a combination of the control students dropping out, taking time off, and delaying. In the first year of college, 88 percent of the control students were enrolled; by the fourth year, control group enrollment had fallen to 78 percent. Persistence in the fourth year is particularly high for those who enrolled at the HI and elite colleges. By shifting students to elite colleges, where they persisted through the first four years more often, the STEM summer programs set the stage for students to graduate from prestigious universities.

B. College Graduation

Only 56 percent of the control students graduated from a four-year college in four years (Figure 2 and Online Appendix Table B.5). This share is higher than the national four-year graduation rate observed at all US institutions of 45.3 percent (U.S. Department of Education 2020) but could be considered low because the students in this sample have near-perfect GPAs and standardized test scores two standard deviations above the countrywide mean. Additionally, the schoolwide four-year graduation rate at the HI for the same cohort is 87 percent. The online program increased the likelihood of graduating with a BA in four years by 3 percentage points, though this difference is not statistically significant. Both the six-week and one-week programs yielded higher graduation rates, but the differences are again not statistically significant. For the six-week and online programs, all of the increases in on-time BA attainment are attributable to graduations from the HI, but the gains from the one-week program are split evenly between the HI and other institutions, according to the conditional randomization estimates. Overall, it appears that the gains in on-time degree attainment are small but that those who received program offers were likelier to graduate from an elite institution on time, with a pooled estimate of 4.8 percentage points (not statistically significant).[xv]

The graduation rates at five- and six-year time horizons are higher for all students, with 77 percent of the control students obtaining a BA within six years after high school graduation

(Figure 2). The one-week and online programs increased BA completion in this longer time window, with particularly pronounced gains at elite institutions for all of the programs. As shown in Table 2, the online program boosted BA attainment by 4 percentage points. Both the pure and conditional randomization estimates show slightly smaller graduation impacts of the six-week program (with a statistically nonsignificant effect of 2–2.5 percentage points) and larger ones for the one-week program (an effect of 9 percentage points). The pooled estimate for graduation from any four-year college within six years is 5.4 percentage points (not statistically significant).

It is notable that the most intensive program delivered the smallest benefit in terms of overall increase in BA attainment within six years. Its weaker effect on BA attainment is most pronounced for elite schools other than the HI. The six-week program increased HI year-one attendance by 17 percentage points and six-year graduation by 14–16 percentage points, indicating little drop-off from attendance to graduation. Among other elite schools, the six-week program increased attendance by 17 percentage points but increased six-year graduation by only 12–13 percentage points. Ultimately, though, the boost to overall BA attainment from the six-week program is not statistically different from that provided by the online group (it is statistically different from that of the one-week group).

The largest shift in graduation patterns was at the HI. All programs increased graduation within six years from the HI, with a pooled estimate of 7.4 percentage points and the largest effects for the six-week program (14–16 percentage points). The one-week program increased graduation within six years at the HI by 6.2 percentage points according to SUR and a not statistically significant 4.5 percentage points according to the conditional randomization estimate. The online program boosted graduation from the HI by 4 percentage points.

Considering institutions beyond the HI, *all* of the programs increased BA attainment in six years from elite colleges. Many control students did graduate from an elite institution in this time frame (43 percent), but we find that, depending on the estimation strategy, the STEM summer programs increased this share by 9–15 percentage points, with a pooled treatment effect of 12 percentage points. Thus, even in the case of the six-week program with its smaller overall graduation boost, students' college trajectories still changed with respect to the type of institutions where they chose to enroll. College quality may enhance social networks and improve human capital, and there is evidence that it increases earnings, especially for URM students (Hoekstra 2009; Dale and Krueger 2002; Dale and Krueger 2014; Zimmerman 2014; Goodman, Hurwitz, and Smith 2017; Chetty et al. 2020; Bleemer 2021).

Overall, assignment to any of the three programs increased enrollment in and graduation from elite colleges, and the one-week and online programs increased BA completion. The gains are concentrated at elite institutions: at the HI for the six-week program and at the HI and other highly ranked colleges for the one-week and online programs. The largest increases in overall graduation came from the one-week program, with the six-week program yielding the smallest gains (the *p*-value for the difference between the six-week and one-week programs is 0.056). The programs succeeded at two of their goals: (1) increasing representation of URM students at the HI and (2) improving the trajectories of students regardless of institution. In the next section, we consider whether the programs induced an increase in attainment of STEM degrees.

C. STEM Degree Attainment and Potential Earnings

One of the programs' stated goals is to increase the proportion of URM students in STEM careers, and a key segment of the STEM pipeline is completion of a degree with a major in a STEM field. We examine STEM degree completion in Column 7 of Table 2. Following the National Center for

Education Statistics, our STEM degree category includes any degree within the broad categorizations used by the Integrated Postsecondary Education Data System (IPEDS) to denote STEM fields: computer and information science, engineering, engineering technologies, biological and biomedical sciences, mathematics and statistics, physical sciences, and science technologies.[xvi] We categorize all other degrees that report a major as non-STEM degrees and also separately report results for degrees with no major code as "missing major" (for details, see Online Appendix Table B.9). Both students who have not achieved the specific degree type (STEM, non-STEM, missing major) and those without degrees are indicated with zeroes in these unconditional outcomes.

In 2019, 36 percent of bachelor's degrees earned in the US were in STEM fields (National Science Board 2022). The study population was predisposed to choosing STEM fields in greater proportion than other US students: 67 percent of the degrees earned within six years in the *control group* were in STEM fields. Even so, the STEM summer programs increased the prevalence of STEM degrees. The online program increased STEM degree attainment by approximately 3 percentage points (not significant). The six- and one-week programs increased STEM degree attainment by approximately 12 percentage points over the figure for the control group (statistically significant). The pooled estimate for any program is 8.5 percentage points (statistically significant at the 10 percent level).

As shown in Online Appendix Table B.9, the six-week program increased overall receipt of four-year degrees by 2–2.5 percentage points: This increase reflects shifts away from non-STEM degrees (of -6 to -7 percentage points), a decrease in the number of unreported majors (of -3 percentage points), and an increase in STEM degree attainment (of 8–11 percentage points).[xvii] The shift to STEM induced by the six-week program operated through the HI. For the one-week program, the STEM gains of 8–10 percentage points are attributable mostly to increased graduation, with a small number of switches away from non-STEM majors (accounting for 2–3 percentage points of the change), though the increases are not statistically significant. For the online program, all of the attainment gains are concentrated in STEM fields at the HI.

Online Appendix Tables B.10 and B.11 separate the STEM majors into specific fields at any institution and the HI, respectively. The tables show that engineering dominated the increase in STEM. Almost all of the students at the HI who were exposed to one of these programs went on to major in engineering. In sum, the programs induced not just college graduation but graduation in STEM fields—particularly engineering.

Differences in STEM degree persistence do not seem to be driven by differences in STEM interest. We turn briefly to our survey evidence: The programs did not change reported interest in STEM majors immediately after the program (Figure 3), with 93 percent of both the treatment and control students at the end of the summer of the program reporting plans to major in a STEM degree. Once in college, both the treatment and control students maintained very high levels of interest in STEM degrees, with approximately 83 percent planning to declare a STEM major. There were some differences in intentions to pursue a STEM career later on: Both the six-week and online programs increased students' likelihood of reporting a desire to pursue a STEM career in the fall of their senior year of high school, and the six- and one-week programs induced similar gains in STEM career intentions mid-college (though the increases are not statistically significant). Thus, we conclude that the increase in STEM degree attainment is due to the groups with program offers being likelier to follow through on their STEM intentions—perhaps because of the upgrade in quality of their institution of enrollment, which we discuss more below—rather than to the programs inspiring a greater degree of interest in STEM fields at (or near) baseline.

To summarize the joint effect of the shifts in choice of college major and college quality, we estimate the programs' impacts on potential earnings. We cannot measure earnings directly, but we use information from the College Scorecard (US Department of Education 2024) as a proxy for the shift in an individual's potential earnings made possible via shifts in major and institution. The College Scorecard reports earnings by institution and major for students who received federal student aid, who may not be representative of the full student body. However, Mabel, Libassi, and Hurwitz (2020) report a high correlation between earnings by major for financial aid recipients and all students. The Scorecard earnings estimates come from both completers and noncompleters and are for students who were a few years older than and entered a different labor market from the one faced by the students in this study. We use median earnings information in 2022 dollars for employed, unenrolled individuals five years after their exit from college in 2014–15 and 2015–16. If information is unavailable for a particular major by institution (typically because of small sample size), we substitute the institution-level outcome for the same cohort. For students who did not attend college, we impute earnings at the fifth percentile of the institution-by-major distribution of earnings (~$36,000). The information is also available by gender, so we report estimates for four relevant outcomes in Table 3: all applicants, where we assign overall median earnings to individuals (Column 1); all applicants with earnings assigned by gender (Column 2); and the same outcomes, but restricted to those who graduated college by their sixth year after high school graduation (Columns 3 and 4). Thus, the first two columns represent unconditional potential earnings, with imputed earnings for nonattendees, and the second two columns condition on college graduation.

The potential earnings of the control group are relatively high, at approximately $100,000. As shown in Panel B of Table 3, regardless of how we compute the outcome, random assignment to the online program increased this by approximately $7,000, or 7 percent. Assignment to the six-week program increased potential earnings to a greater extent, and the one-week program had a smaller effect than the online one, though the difference in both cases is not statistically significant (Panel A). Our summary measures of these effects in Panels C and D show, depending on the specification, that the six-week program raised potential annual earnings by $11,000–$15,000 and the one-week program boosted earnings by $3,000–$5,000, though this latter difference is not statistically significant. The pooled estimate is approximately $7,000–$8,000. In most models, the estimates for the six-week program are higher than the estimates for the one-week and online programs (see *p*-values at the bottom of Panel D). This is because of the increase in the number of students choosing highly remunerative engineering majors concentrated at the HI. Though the differences between the one-week and online programs are not statistically significant, it is notable that the one-week program had smaller potential earnings effects than the online program when it appears to have had larger graduation effects than the online program. This is because the earnings estimates operate through the institution–major–gender level and the online program prompted a large shift into engineering and affected more men. Of course, the College Scorecard measures for earlier cohorts are not exact earning measures, but they provide a summary measure indicating that the programs induced students to embark on college and major trajectories with the potential to increase their earnings by 3–15 percent.[xviii]

D. Impacts for Subgroups

In Online Appendix Tables D.1–D.8, we present estimates for key outcomes by gender, URM status, self-reported free or reduced-price lunch receipt (a proxy for family income), and first-generation college-going status. These groups correspond to populations of particular interest for increasing representation in STEM fields. However, we note that estimating effects for these subgroups splits our small sample, reducing precision. Thus, we discuss here only the differences between subgroups that are statistically significant. It does appear that the one-week program raised women's likelihood of graduating with a STEM degree more than it raised men's (p-value = 0.021). This may be because women are less likely to choose STEM majors so there is more room for improvement. Thus, STEM programs that also emphasize serving women (by ensuring gender balance in the program, even while not explicitly focusing on women's representation) can also contribute to closing gender gaps in STEM. However, this is just one outcome and one program among many, and overall, both men and women benefited from the STEM summer programs.

In Online Appendix Tables D.3 and D.4, we present separate estimates for Black, Hispanic, and Native American individuals and for students who do not identify themselves as such (white, Asian, multiethnic, and "other race" students are in the non-URM category). Because less than 18 percent of the sample is non-URM, these estimates are noisy, but there is some evidence, at least for STEM degree attainment, that the non-URM impacts are larger for the six- and one-week programs. However, it appears that the benefits of the online program accrued mostly to URM students. Again, however, the sample of non-URM students is quite small, and there are meaningful benefits for both groups.

We observe the clearest difference between students who reported receiving subsidized lunch at their high school and those who did not (Online Appendix Tables D.5 and D.6). Although exposure to the programs was generally beneficial for both groups, students who did *not* receive subsidized lunch reaped larger gains, especially in regard to graduating from an elite institution and graduating with a STEM degree; for some outcomes, the effect for subsidized lunch students was a noisy zero. Students with more resources may be better poised to upgrade their choice of enrollment institution, perhaps because of fewer concerns about college costs or ease of the social transition. This also implies that increasing representation by race in STEM fields will not necessarily increase representation by family income and vice versa. The differences by first-generation college-going status (defined as neither parent having ever attended a four-year college) align with the findings by subsidized lunch status but are generally more muted (Online Appendix Tables D.7 and D.8).

E. Robustness

We consider how robust our estimates are to alternative specifications and whether our conclusions remain similar under randomization inference. We also discuss the role of treatment effect heterogeneity.

1. Alternative specifications

Online Appendix Table C.1 shows the main estimates without baseline covariates. As expected, precision decreases somewhat, but the magnitude and sign of the estimates remain the same. Online Appendix Tables C.2–C.4 remove each cohort in turn. We expect there to be idiosyncrasies across cohorts due to sampling variation, but each table reports estimates that are generally in line with the main findings. In some cases, precision declines because the samples are smaller. Online Appendix Table C.2 is notable in that it removes the 2014 cohort. This cohort was subject to the most modifications to random assignment, featuring more blocks than the two subsequent cohorts. We would expect that if the study design with the greater number of blocks did not remove selection bias as thoroughly as the design with fewer blocks, the conditional randomization estimates in this table should be *smaller* than the main estimates because our removing the observations from 2014 should reduce upward bias in the estimates. Instead, the results limited to the later cohorts with a design that more closely approximates completely random assignment are very similar.

2. Randomization inference

A different threat to validity is associated with the randomization scheme. By imposing randomization strata and modifying complete random assignment, the study limits the possible range of observable outcomes. For example, if a state is preferred in the assignment for representational reasons, this preference will always limit the range of randomization scenarios possible under the assignment scheme. Standard inference methods do not account for these constraints (Athey and Imbens 2017). As an alternative, we present results derived by randomization inference. Specifically, we rerandomize applicants into the programs using the same randomization criteria (blocks, location, gender preferences, etc.) 1,000 times and compare the estimate from our main specification with the distribution of estimates from the 1,000 randomizations. Each randomization faces the constraints imposed by our research design, so we effectively limit the possible comparisons to assignments that might actually have occurred rather than those that could have been made in a hypothetical scenario with full randomization. We conduct this exercise for both the pure and conditional randomization estimates, as both of these approaches are subject to at least some constraints.

We display the results from this exercise in Online Appendix Figures C.1 and C.6: first the coefficients from pure randomization (with the six- and one-week programs compared to the online program and the online group compared to the control group) and then the coefficients from conditional randomization (all treatments compared to the control). Each panel shows the distribution of treatment estimates from the 1,000 hypothetical randomizations (bars), compared to the relevant estimate (dashed line). If the treatment effect estimate from the tables is at or above the 97.5th percentile (two-tailed test) or 95th percentile (one-tailed test), this implies statistical significance.

For pure randomization, the impact estimates for the online program vs. the control group are often at the 99th percentile, and when they are lower, it is for the outcomes that are also not statistically significant under standard inference methods (attendance of and graduation from any four-year institution, STEM degree attainment). The impact estimates for the six- and one-week programs vs. the online group are never above the 95th percentile, but again, this parallels our results from traditional inference methods, where the Block 1 comparisons tend to have less statistical power. For conditional randomization, the impact estimates for attendance of any four-year institution fall at the 63rd–82nd percentiles, but for elite institutions, the impacts on attendance are at the 98th–99th percentiles—generally similar to the pattern of findings from the main estimates. The impacts on elite college graduation fall within the 96th–98th percentiles. Again, this pattern aligns with the trend of our main estimates. The impacts on STEM degree completion follow the pattern in the estimates from traditional inference, with statistically significant effects for the six- and one-week programs but not for the online program. We thus consider the randomization inference exercise to offer confirmatory evidence that we can draw inferences from both of our randomization schemes using traditional statistical methods.

3. Testing constant treatment effects

Another threat to validity would arise if the constant treatment effects assumption that we discussed in Section IV.C failed to hold. The SUR and conditional randomization estimation strategies rely on the effect of the online program being the same in Blocks 1 and 2, as shown in Section IV.C. If treatment effects were not constant, we could not make the cancellation of online treatment effects across blocks that made extrapolation possible, and instead there would be a bias term in our estimate of the six- and one-week program effects, as can be seen here:

$$
\begin{aligned}
\delta_{6w} &= \\
&= E[\,Y_{6w} \mid Block1, R_k\,] - E[\,Y_{on} \mid Block1, R_k\,] + E[\,Y_{on} \mid Block2, R_k\,] - E[\,Y_0 \mid Block2, R_k\,] \\
&= \underbrace{E[\,Y_{6w} \mid Block1, R_k\,] - E[\,Y_0 \mid Block2, R_k\,]}_{ATE_{6w}} + \underbrace{E[\,Y_{on} \mid Block2, R_k\,] - E[\,Y_{on} \mid Block1, R_k\,]}_{Bias}.
\end{aligned}
$$

We assess the existence of bias in Online Appendix Table C.5, which displays the regression-adjusted differences in outcomes in the online assigned group between Blocks 2 and 1 (the bias term above). Panel A displays the difference controlling for the covariates we include in all of our other estimates and for randomization strata that exclude block assignment. Panel B adds a control for the rating variable, which is used to assign students to blocks. There are mostly small, not statistically significant differences between the two online groups, especially in Panel A. In Panel B, which controls for rating variable as a substitute for block, the differences are little larger, but only statistically significant for graduation from an elite college within six years, with an 8-percentage-point difference. If we take that 8-percentage-point estimate as a given, that implies the six- and one-week programs did not yield the 13- to 14-percentage-point increase in graduation from elite schools vs. the control group and instead had a 5- to 6-percentage-point effect. Nevertheless, the pure random assignment estimate for the online program still indicates a 9-percentage-point increase in graduation from these schools (Panel B of Table 2), and this is the only outcome for which we saw a statistically significant difference in performance level within the online program by block. Additionally, the constant treatment effects concern affects the interpretation only of the SUR estimates (Panel C) and conditional randomization estimates (Panel D) because these are the estimates that involve extrapolating across blocks.

VI. Mechanisms

In this section, we explore mechanisms behind the success of the STEM summer programs, focusing on graduation from a four-year college within six years and STEM degree attainment as the main outcomes of interest. For many of these analyses, we use data from the student surveys, which are subject to some differential response by treatment arm. As we noted in Section III, we use the unweighted survey responses because of the inconsistent influence of reweighting schemes in addressing survey response bias, and we instead consider the findings from the surveys suggestive rather than conclusive.

There are several reasons why the programs might have improved college graduation and STEM degree attainment. We focus on three human capital–related explanations here. First, the program may have increased participants' human capital with respect to the college application process, which in turn could have enhanced the quality of the institution chosen for enrollment and given rise to the differences in outcomes that we observe. We also discuss the extent to which such institutional upgrading could be due to greater likelihood of admission because of a signaling effect of the programs. Second, the program may have improved participants' human capital with respect to subject matter knowledge, helping them get ahead in college and be likelier to graduate. Finally, we consider whether there was an increase in participants' human capital with respect to their soft skills, which could have better positioned them to succeed in college. As we detail below, the evidence that the first channel explains our findings is strongest, though we cannot fully rule out the other explanations.

A. College Application Behavior and Shifts in Quality of College of Enrollment

The STEM summer programs have the potential to affect young people's trajectories, but ultimately, even the most intense six-week program covers only a short period in a young person's life. However, by prompting students to apply to better colleges than they would have considered otherwise, the programs may have set students on a different path. We have shown that program participation increased enrollment in and graduation from the HI and other highly ranked colleges. Here, we show that the primary path whereby students shifted into these institutions was through a change to their college application strategy. We first show that application and admissions patterns changed for program participants and then discuss how this shift induced students to enroll in higher-quality institutions. In administrative data from the HI, we observe increased applications to the HI, as displayed in Online Appendix Table B.8, and survey data shows a rise in applications to other elite institutions, as displayed in Table 4. Approximately 32 percent of the control students applied to the HI. This would be a very high rate of application for typical high school students, but the program applicants are a selected group who were interested in the outreach office's summer programs. Even so, assignment to any of the summer programs more than doubles the rate of application to the HI, with almost all students in the six-week group (78 percent) applying to the HI.

Almost 6 percent of the control students were admitted to the HI (approximately 19 percent of those who applied), again demonstrating that the study sample is a selected group: The HI typically admits fewer than 10 percent of applicants. Admission rates are even higher among those with program offers. According to the pure estimates, approximately 37 percent of students assigned to the six-week program were admitted to the HI, with an implied admission rate of 47 percent. For the one-week program, approximately 29 percent were admitted, with an implied admission rate of about 38 percent; for the online program, 15 percent were admitted, with an implied admission rate of approximately 22 percent. Thus, the six-week and one-week programs delivered a bump in the likelihood of admission; admissions rates are similar between the control group and the online group. Greater admission to the HI therefore seems to be due to two factors:

the students' greater likelihood of applying to the HI, and for those in the six- and one-week programs, their greater likelihood of being admitted to it. We explore the potential of a signaling effect contributing to this in more detail below.

Enrollment and graduation changed beyond the HI, as well, with the programs, especially the one-week and online programs, also shifting students toward elite institutions other than the HI, so admission behavior with respect to these institutions likely changed, too. We turn to the survey responses to examine changes in application behavior more broadly. Survey respondents reported the colleges they applied and were admitted to in a survey conducted in May of their senior year of high school. As we show in Table 4, those offered a seat at a STEM summer program were likelier to have applied and been admitted to elite institutions, even when the HI is excluded from that category (Column 5). These differences are statistically significant for the six-week and online programs, but not for the one-week group. The programs also induced an increase in the overall number of applications and admissions (Column 2), though only the online program's effect on applications is statistically significant. However, the most dramatic change in application behavior is the reduction in applications to a single school (Column 4). Attending the six-week program eliminated the (small) possibility of students applying to only one school, and attendance of the one-week or online program greatly diminished this likelihood. The declines for the six-week and online programs are statistically significant, but that for the one-week program is not.

These improvements in college application behavior likely arose from increased knowledge about the college admissions process. The programs also improved students' college resources by enhancing their knowledge of both the landscape of available colleges and the college application process itself, as we show in Online Appendix Table E.3. The survey responses from the fall of the senior year of high school show that the programs—especially the six-week program, for which we most consistently observe statistically significant differences—increased sources of college application advice (Columns 1–4), with students assigned to a treatment group reporting a greater likelihood of obtaining advice from friends and a teacher or counselor. Assignment to a program also increased familiarity with non-HI institutions: Students were likelier to report familiarity with a technical institution (by 7–10 percentage points), an elite institution (by 1 percentage point), and a liberal arts college (by 7–10 percentage points) and were less likely to report familiarity with a highly ranked public university or a fake institution with a made-up name, though these differences are not significant.

These changes in knowledge of the college application process and application behavior translated into increased enrollment in high-quality institutions, as we show in Section V. We investigate whether enrollment in these institutions spurred the college graduation and STEM results we observe in Figure 4, which compares the *institution-level* outcome of the university attended immediately after high school graduation to the individual-level outcome. To generate the institution-level outcomes, we assign the institution-specific graduation rate to students who attended a four-year institution in their first year of college.[xix] This outcome then measures the "predicted" graduation rates for each treatment group, assuming that the students in the study group graduated at the four- and six-year graduation rate of the institution they attended. We construct a similar outcome for STEM degree attainment, namely, the institution-specific share of degrees in STEM fields, on the basis of degrees reported to IPEDS. This outcome does not differentiate by time to graduation, so it is the same for the comparisons of graduation by year four and by year six.

The results from this exercise appear in Figure 4, with more details in Online Appendix Table B.15. Panel A compares the difference in "actual" graduation with the "predicted" graduation rates for graduation within four years and Panel C for graduation within six years. For the four-year graduation rates, the change in actual graduation closely parallels the difference in predicted graduation rates (for the six-week and online programs, the change in graduation is slightly smaller than the difference in expected graduation), consistent with many of the program benefits operating through upgrades to students' college of enrollment. For the six-year graduation rates, the actual graduation rate almost exactly matches the predicted graduation rate for the online program, and the actual rate exceeds the predicted one for the one-week program. For the six-week program, the actual graduation boost is approximately half of what we would expect given the prediction.

When we conduct a similar analysis for the STEM degree share, the results are noisier. We see in Panels B and D that the six-week and online programs increased attendance at institutions with higher proportions of STEM degrees but the one-week program did not. The gains in STEM degree receipt for the one-week program are larger than the (very small) predicted increase. The gains in actual STEM degree receipt are larger than the predicted difference for the other two programs, but the confidence intervals overlap. Thus, students offered STEM summer programs generally had a higher likelihood of graduating because of the upgrades to their choice of institution of enrollment, but they may have increased their STEM degree attainment to an even greater extent than predicted, which provides evidence that program participation helped students achieve their intention of obtaining STEM degrees in very competitive institutions. The evidence on college quality is consistent with the finding of Barr and Castleman (2025) that approximately 70 percent of the gains in college graduation due to an intensive advising program came from upgrades in institutional quality.

B. Signaling

We next consider that participants may have signaled their quality by mentioning their program participation in their college applications as a potential explanation, though it is difficult to disentangle the operation of a signaling mechanism from that of a human capital mechanism. Given that the programs were hosted at the HI, any influence of signaling would probably have proved most powerful there. Because many of the gains from the one-week and online programs were reaped at non-HI elite institutions, the possibility for signaling to play a large role in explaining those effects is limited. Additionally, although signaling may have had an impact, we note that it would primarily have influenced *admission*, not necessarily what happened during college. Thus, it is unlikely to explain the increase in STEM degree attainment, especially given that students in both the treatment and control groups reported an intention to major in STEM fields at very high rates.

We examine how much of the admission effect is due to more ambitious application behavior as opposed to greater odds of admission upon college application in Figure 5 and Online Appendix Table B.14 by using two comparison groups. In this figure, we show application, admission, and admission conditional on applying to the HI (Panels A, C, and E) and to any elite institution excluding the HI (Panels B, D, and F). The results in the latter group of panels are restricted to the survey respondents who reported where they were admitted and where they enrolled. The conditional admission estimates show the increase in admission to an institution for those offered a place in the STEM summer programs, an effect that may be due to signaling,[xx] program-induced improvements in students' application packages or human capital, or a change in the type of applicant. Part of the admissions gains at the HI are attributable simply to the increased likelihood of application, but a share of the gains are from the additional bump in admissions likelihood from attendance of the STEM summer programs (14 percentage points for the six-week program and 5 percentage points for the one-week and online programs, though the latter two differences are not statistically significant). At non-HI elite institutions, the online and one-week programs increased the likelihood of conditional admission by approximately 5 percentage points. Overall, these findings show that there were some increases in conditional admission but that the role of changes in application behavior is large.

A second comparison in this figure compares the students in the HI STEM summer programs to a control group limited to students who reported having *any* STEM summer experience.[xxi] Students in this control group are likelier to be female and have slightly higher rating scores than the control students without STEM summer experiences, but they are otherwise quite similar to the control group overall (see Online Appendix Table B.17). If we assume that different STEM summer experiences impart similar human capital gains, the comparison of those assigned to the HI STEM summer programs with those in the control group with STEM summer experiences tests whether the HI programs had signaling value above and beyond that of involvement in any STEM summer programming.

Even with this comparison group, there is still a conditional admission gain of 10 percentage points at the HI for attendees of the six-week program, but there is no difference for participants of the one-week and online programs. We interpret this result to mean that there is likely some role of signaling in explaining the effects of the six-week program at the HI. However, outside the HI, there appears to be little difference in conditional admission between the HI STEM summer groups and control group members with other STEM summer experiences, implying that, at other institutions, there were no advantages in conditional admission from the summer programs we study here relative to those of any STEM summer programs, except the gains delivered through the application channel. These estimates are consistent with some role for signaling but do not prove its existence. Even when we compare treated students to control students with STEM summer experiences, our assumption of similar human capital gains across programs could be wrong, and the additional boost for students in the six-week program relative to control students with STEM summer experiences might be due to human capital gains from the program. Alternatively, the six-week program may have changed the type of applicant, even compared to other STEM summer controls, but the one-week and online programs did not. These findings set a ceiling on the potential effects operating through the signaling mechanism and confirm that the application channel accounts for much of the programs' effects.

C. Subject Matter Knowledge, Skills, and Confidence

Another way the STEM summer programs may have boosted college graduation and STEM degree attainment is by increasing students' STEM subject matter knowledge, helping them get ahead in college. This could have occurred in two ways: The programs could have changed which high school classes the students took prior to college entrance, or they could have directly increased the students' knowledge of STEM, better preparing them for future STEM majors. Additionally, the programs may have improved more general skills such as study skills and soft skills such as confidence and self-esteem.

Table 5 shows impacts on both high school course plans and a direct measure of human capital. Most students were already planning to take at least one AP or IB course in their senior year in high school (the control group mean is 92 percent), and program participation had positive impacts on these intentions, as shown in Column 1. The programs also increased the likelihood of students taking at least one AP or IB course by 3–12 percent, with statistically significant differences for the one-week and online groups. Much of this increase flowed through computer science take-up, with a statistically significant difference in AP Computer Science for the six-week program (and small declines in other AP courses) and statistically nonsignificant gains for the one-week and online programs. Approximately 14 percent of the control group planned to enroll in such a course, and receipt of a program offer increased this share by 7–10 percentage points, though the difference is statistically significant only for the six-week program. Perhaps because science and math advanced coursework was already extremely popular (with control means of approximately 75 percent), the STEM programs made little difference for advanced science and math coursework. During the fall of their senior year of high school, applicants offered seats in the programs were slightly abler to answer a calculus question, though the differences are not statistically significant (Column 5). The minor calculus improvement and coursework changes in noncore subjects indicate that, although the programs may have imparted STEM knowledge and inspired subsequent high school coursework, these changes are likely to have been relatively small.

We also illustrate the program impacts on soft skills in Table 5, which shows student responses to survey questions about life skills, study skills, confidence, interest in learning, and attention span.[xxii] Life skills include tasks such as setting an alarm to be on time and doing one's laundry. For many program participants, the time on the HI campus might have been their first time away from their family, and we indeed observe an increase in the self-reported life skills that one would gain in such a situation. Unsurprisingly, the gains in life skills are larger for the residential programs than for the online program, even though this program had a short campus visit component. We also see that the programs increased self-reported study skills, such as asking questions and taking notes. However, there is no statistically significant change in confidence, including students' self-assessed math ability. The positive, statistically significant increases in life and study skills are all consistent with the idea that the STEM summer programs better prepared students to succeed in and graduate from college and may have contributed to the increases in STEM degree attainment that we cannot account for solely through the upgrades to students' college of enrollment.

VII. Conclusion

We present evidence from the first RCT investigating the impact of STEM summer programs for underrepresented youth, examining a suite of programs that includes a six-week, a one-week, and an online program. Receipt of a program offer increased student matriculation at the HI and other elite universities. The programs increased graduation within six years from the HI and other elite institutions, with the one-week and online programs increasing overall BA receipt. The programs also induced increases in attainment of STEM degrees, particularly in engineering. The shift in institutional quality and major may have increased potential earnings by 3–15 percent. We show evidence consistent with the idea that a change in college application behavior shifted students into higher-quality institutions, which drove the gains in college graduation. For the six-week program, the gains in institutional quality primarily occurred because of enrollment in the HI, but for the one-week and online programs, they were split between the HI and other elite institutions. Although these programs occurred at a particular elite intuition and may not be replicable elsewhere, our findings highlight the important roles of college application and college quality, which are applicable in multiple college access contexts.

Much of the gains in STEM degree completion came from increases in bachelor's degree attainment for this highly STEM-motivated group, but some of the gains came from the programs preventing switches out of STEM majors. Almost all of the study sample intended to major in STEM in college, but such intentions often flounder during college, giving rise to the "leaky pipeline" problem, whereby students leave STEM tracks throughout college. The STEM summer programs made students more resistant to pipeline leaks, though there were still STEM-interested students who did not ultimately complete a STEM degree. We cannot definitively say why this occurred, but we note that the programs offered comprehensive coverage of many hypothesized STEM-supportive pathways, including STEM curricula; URM role models in the form of near-peer teaching and residential assistants, staff, instructors, and guest speakers; and a shared group experience.

The contrast between the three programs provides some basis for a back-of-the-envelope cost-effectiveness calculation. The six-week and one-week programs generated similar responses, though differing in their intensity. For STEM degree completion, the online program had a smaller effect, but it increased graduation from elite institutions almost as much as the residential programs and actually had a larger overall graduation effect than the six-week program. In 2015, the six-week program cost approximately $15,000 per student; the one-week and online programs cost approximately $2,000 per student. Perspectives on which program is the most cost-effective will differ by the objective. For the HI, the interest may be greatest in HI graduation, the outcome on which the six-week program had a large impact—though this is in part because of the higher admission rates for this group. However, if a policymaker is interested in supporting overall graduation from elite institutions, the six-week program underperformed relative to the one-week program by a small amount, though its impacts were larger than those of the online program. The residential programs did outperform the online program in terms of STEM degree completion (Cohodes et al. 2024).

Simply comparing the program costs to the potential earnings estimates in Table 3 provides back-of-the-envelope estimates of the cost-effectiveness of the program. The increase in one year of potential earnings at five years after graduation is approximately $11,000–$15,000 for the six-week ($15,000 cost) program, $3,000–$6,000 for the one-week ($2,000 cost) program, and $6,000–$7,400 for the online ($2,000 cost program). The estimates used here do have large standard errors; for example, the earnings gains for the online program may be as large as $12,000 or as small as $2,000. A single donation to the HI of a years' gain in potential earnings as measured by the point estimates would almost (in the case of the six-week program) or more than (in the case of the one-week and online programs) cover the institution's costs, though the programs are not funded solely via donations. If the provider of a similar program were the government, it could certainly expect returns from increased tax revenues due to higher earnings. The earnings gains would likely accumulate over a lifetime of earnings. Thus, even if the true impacts on earnings are toward the lower end of the confidence intervals, these gains will exceed the costs of the program over time.

This analysis shows that targeted programs to increase representation on college campuses can have wide-ranging benefits for participants. There may be additional spillover benefits for peers at elite institutions, who can benefit from a more diverse and inclusive STEM classroom. With the 2023 US Supreme Court decision terminating the use of affirmative action in college admissions, more colleges and universities may turn to programs such as the STEM summer programs we study here to provide the benefits of diversity to their campuses through indirect avenues. Indeed, many campuses already have "summer bridge" programs that provide support for matriculating underrepresented students in the summer before their freshman year. Additionally, federal investment in STEM fields is targeted at higher education rather than earlier segments of the pipeline. Our findings show that a focus on higher education *after* students apply to college may neglect a key opportunity to intervene in students' lives *before* they apply to college—a crucial point in time when students choose the institutional affiliations that may ultimately boost their success.

References

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel. 2016. "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." In *Handbook of the Economics of Education*. Vol. 5, ed. Eric A. Hanushek, Stephen Machin and Ludger Woessmann. Elsevier.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers." *Annual Review of Economics*, 4: 185–223.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103: 1481–1495.

Andrews, Rodney J, Scott A Imberman, and Michael F Lovenheim. 2020. "Recruiting and Supporting Low-Income, High-Achieving Students at Flagship Universities." *Economics of Education Review*, 74: 101923.

Arcidiacono, Peter. 2004. "Ability Sorting and the Returns to College Major." *Journal of Econometrics*, 121(1-2): 343–375.

Arcidiacono, Peter, and Michael Lovenheim. 2016. "Affirmative Action and the Quality-Fit Trade-Off." *Journal of Economic Literature*, 54(1): 3–51.

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economics Review*, 106(3): 525–562.

Astorne-Figari, Carmen, and Jamin D Speer. 2019. "Are Changes of Major Major Changes? The Roles of Grades, Gender, and Preferences in College Major Switching." *Economics of Education Review*, 70: 75–93.

Athey, Susan, and Guido W Imbens. 2017. "The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*. Vol. 1, 73–140. Elsevier.

Avery, Christopher. 2010. "The Effect of College Counseling on High-Achieving, Low-Income Students." National Bureau of Economic Research Working Paper 16359.

Avery, Christopher. 2013. "Evaluation of the College Possible Program: Results from a Randomized Controlled Trial." National Bureau of Economic Research Working Paper 19562.

Bailey, Martha J, and Susan M Dynarski. 2011. "Gains and Gaps: Changing Inequality in US College Entry and Completion." National Bureau of Economic Research.

Baker, Rachel, Daniel Klasik, and Sean F Reardon. 2018. "Race and Stratification in College Enrollment over Time." *AERA Open*, 4(1): 2332858417751896.

Barr, Andrew, and Benjamin Castleman. 2025. "Increasing Degree Attainment among Low-Income Students: The Role of Intensive Advising and College Quality." *American Economic Review*, 115(11): 4075ˆa4103.

Becker, Charles M., Cecilia Elena Rouse, and Mingyu Chen. 2016. "Can a Summer Make a Difference? The Impact of the American Economic Association Summer Program on Minority Student Outcomes." *Economics of Education Review*, 53: 46–71.

Bettinger, Eric. 2010. "To Be Or Not To Be: Major Choices in Budding Scientists." In *American universities in a global market*. 69–98. University of Chicago Press.

Bettinger, Eric P, and Bridget Terry Long. 2005. "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." *American Economic Review*, 95(2): 152–157.

Bleemer, Zachary. 2021. "Top Percent Policies and the Return to Postsecondary Selectivity."

Bleemer, Zachary. 2022. "Affirmative Action, Mismatch, and Economic Mobility After California's Proposition 209." *The Quarterly Journal of Economics*, 137(1): 115–160.

Bradford, Brittany C, Margaret E Beier, and Frederick L Oswald. 2021. "A Meta-Analysis of University STEM Summer Bridge Program Effectiveness." *CBE Life Sciences Education*, 20(2): ar21.

Budig, Michelle J., Misun Lim, and Melissa J. Hodges. 2021. "Racial and Gender Pay Disparities: The Role of Education." *Social Science Research*, 98: 102580.

Carrell, Scott, and Bruce Sacerdote. 2017. "Why Do College-Going Interventions Work?" *American Economic Journal: Applied Economics*, 9(3): 124–51.

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly journal of economics*, 125(3): 1101–1144.

Castleman, Benjamin, and Joshua Goodman. 2018. "Intensive College Counseling and the Enrollment and Persistence of Low-Income Students." *Education Finance and Policy*, 13(1): 19–41.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

Cohodes, Sarah, Helen Ho, and Silvia Robles. 2022. "A Randomized Evaluation of STEM Focused Summer Programs." AEA RCT Registry, Number: AEARCTR-0002888. DOI: https://doi.org/10.1257/rct.2888.

Cohodes, Sarah R., and Joshua S. Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy." *American Economic Journal: Applied Economics*, 6(4): 251–85.

Cohodes, Sarah R., Helen Ho, Elizabeth Huffaker, and Silvia C. Robles. 2024. "Residential versus Online? Experimental Evidence on Diversifying the STEM Pipeline." *AEA Papers and Proceedings*, 114: 507ˆa11.

Cook, Lisa D., Janet Gerson, and Jennifer Kuan. 2022. "Closing the Innovation Gap in Pink and Black." *Entrepreneurship and Innovation Policy and the Economy* 1.1 (2022): 43-66.

Cosentino, Clemencia, Cecilia Speroni, Margaret Sullivan, Raul Torres, et al. 2015. "Impact Evaluation of the RWJF Summer Medical and Dental Education Program (SMDEP)." Mathematica Policy Research.

Dale, Stacy, and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics*, 117(4): 1491–1528.

Dale, Stacy, and Alan B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources*, 49(2): 323–358.

Darolia, Rajeev, Cory Koedel, Joyce B Main, J Felix Ndashimye, and Junpeng Yan. 2020. "High School Course Access and Postsecondary STEM Enrollment and Attainment." *Educational Evaluation and Policy Analysis*, 42(1): 22–45.

De Philippis, Marta. 2021. "STEM Graduates and Secondary School Curriculum: Does Early Exposure to Science Matter?" *Journal of Human Resources*, 1219–10624R1.

Dillon, Eleanor Wiske, and Jeffrey Andrew Smith. 2017. "Determinants of the Match Between Student Ability and College Quality." *Journal of Labor Economics*, 35(1): 45–66.

Dutz, Deniz, Ingrid Huitfeldt, Santiago Lacouture, Magne Mogstad, Alexander Torgovitsky, and Winnie van Dijk. 2021. "Selection in Surveys." National Bureau of Economic Research Working Paper 29549.

Dynarski, Susan, CJ Libassi, Katherine Michelmore, and Stephanie Owen. 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): 1721–56.

Fairlie, Robert W, Florian Hoffmann, and Philip Oreopoulos. 2014. "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." *American Economic Review*, 104(8): 2567–91.

Fischer, Stefanie. 2017. "The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence." *Labour Economics*, 46: 211–226.

Gerber, Theodore P, and Sin Yi Cheung. 2008. "Horizontal Stratification in Postsecondary Education: Forms, Explanations, and Implications." *Annual. Review of Sociology*, 34: 299–318.

Goodman, Joshua, Michael Hurwitz, and Jonathan Smith. 2017. "Access to 4-Year Public Colleges and Degree Completion." *Journal of Labor Economics*, 35(3): 829–867.

Granovskiy, Boris. 2018. "Science, Technology, Engineering, and Mathematics (STEM) Education: An Overview. CRS Report R45223, Version 4." *Congressional Research Service*.

Green, Alan, and Danielle Sanderson. 2018. "The Roots of STEM Achievement: An Analysis of Persistence and Attainment in STEM Majors." *The American Economist*, 63(1): 79–93.

Griffith, Amanda L. 2010. "Persistence of Women and Minorities in STEM Field Majors: Is it the School that Matters?" *Economics of Education Review*, 29(6): 911–22.

Griffith, Amanda L, and Joyce B Main. 2019. "First Impressions in the Classroom: How Do Class Characteristics Affect Student Grades and Majors?" *Economics of Education Review*, 69: 125–137.

Hastings, Justine S., Christopher A. Neilson, and Seth D. Zimmerman. 2013. "Are Some Degrees Worth More Than Others? Evidence from College Admission Cutoffs in Chile." National Bureau of Economic Research Working Paper 19241.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *The Review of Economics and Statistics*, 91(4): 717–724.

Hoffmann, Florian, and Philip Oreopoulos. 2009. "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources*, 44(2): 479–494.

Hofstra, Bas, Vivek V Kulkarni, Sebastian Munoz-Najar Galvez, Bryan He, Dan Jurafsky, and Daniel A McFarland. 2020. "The Diversity–Innovation Paradox in Science." *Proceedings of the National Academy of Sciences*, 117(17): 9284–9291.

Hoxby, Caroline, and Christopher Avery. 2013. "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students." *Brookings Papers on Economic Activity*, 1–65.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow. 2019. “The Allocation of Talent and US Economic Growth.” *Econometrica*, 87(5): 1439–1474.

Joensen, Juanna Schrøter, and Helena Skyt Nielsen. 2016. “Mathematics and Gender: Heterogeneity in Causes and Consequences.” *The Economic Journal*, 126(593): 1129–1163.

Kaganovich, Michael, Morgan Taylor, and Ruli Xiao. 2023. "Gender differences in Persistence in a Field of Study: This Isn’t All About Grades." *Journal of Human Capital* 17(4): 503-556.

Kinsler, Josh, and Ronni Pavan. 2015. “The Specificity of General Human Capital: Evidence from College Major Choice.” *Journal of Labor Economics*, 33(4): 933–972.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. “Field of Study, Earnings, and Self-Selection.” *The Quarterly Journal of Economics*, 131(3): 1057–1111.

Kitchen, Joseph A, Gerhard Sonnert, and Philip M Sadler. 2018. “The Impact of College-and University-Run High School Summer Programs on Students’ End of High School STEM Career Aspirations.” *Science Education*, 102(3): 529–547.

Kitchen, Joseph A, Philip Sadler, and Gerhard Sonnert. 2018. “The Impact of Summer Bridge Programs on College Students’ STEM Career Aspirations.” *Journal of College Student Development*, 59(6): 698–715.

Kugler, Adriana D, Catherine H Tinsley, and Olga Ukhaneva. 2021. “Choice of Majors: Are Women Really Different from Men?” *Economics of Education Review*, 81: 102079.

Lovenheim, Michael, and Jonathan Smith. 2023. "Returns to Different Postsecondary Investments: Institution Type, Academic Programs, and Credentials." In *Handbook of the Economics of Education*, 6: 187-318. Elsevier.

Mabel, Zachary, CJ Libassi, and Michael Hurwitz. 2020. “The Value of Using Early-Career Earnings Data in the College Scorecard to Guide College Choices.” *Economics of Education Review*, 75: 101958.

National Science Board. 2021. “The STEM Labor Force of Today: Scientists, Engineers and Skilled Technical Workers: Science and Engineering Indicators 2021.” National Science Foundation NSB-2021-2, Arlington, VA.

National Science Board. 2022. “Higher Education in Science and Engineering: Science and Engineering Indicators 2022.” National Science Foundation NSB-2022-3, Arlington, VA.

Nomi, Takako, Darrin DeChane, and Michael Podgursky. 2025. "Project Lead the Way: Impacts of a High School Applied STEM Program on Early Post-Secondary Outcomes." *Journal of Research on Educational Effectiveness*: 1-29.

Owen, Stephanie. 2023. "College Major Choice and Beliefs about Relative Performance: An Experimental Intervention to Understand Gender Gaps in STEM." *Economics of Education Review* 97 (2023): 102479.

Owen, Stephanie. 2021. "Ahead of the Curve: Grade Signals, Gender, and College Major Choice."

Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2014. "The Nexus Between Labor Diversity and Firm's Innovation." *Journal of Population Economics*, 27(2): 303–364.

Price, Gregory N. 2005. "The Causal Effects of Participation in the American Economic Association Summer Minority Program." *Southern Economic Journal*, 72(1): 78–97.

Price, Joshua. 2010. "The Effect of Instructor Race and Gender on Student Persistence in STEM Fields." *Economics of Education Review*, 29(6): 901–910.

Riegle-Crumb, Catherine, Barbara King, and Yasmiyn Irizarry. 2019. "Does STEM Stand Out? Examining Racial/Ethnic Gaps in Persistence Across Postsecondary Fields." *Educational Researcher*, 48(3): 133–144.

Robles, Silvia. 2018. "The Impact of a STEM-Focused Summer Program on College and Major Choices Among Underserved High-Achievers." Blueprint Labs Discussion Paper 2018.03.

Sass, Tim R. 2015. *Understanding the STEM Pipeline. Working Paper 125.* National Center for Analysis of Longitudinal Data in Education Research.

U.S. Department of Education. 2024. "College Scorecard Data [Dataset]." U.S. Department of Education.

U.S. Department of Education. 2020. "Digest of Education Statistics 2020. Table 326.10." *National Center for Education Statistics*.

Yang, Yang, Tanya Y. Tian, Teresa K. Woodruff, Benjamin F. Jones, and Brian Uzzi. 2022. "Gender-Diverse Teams Produce More Novel and Higher-Impact Scientific Ideas." *Proceedings of the National Academy of Sciences*, 119(36): e2200841119.

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics*, 32(4): 711–754.

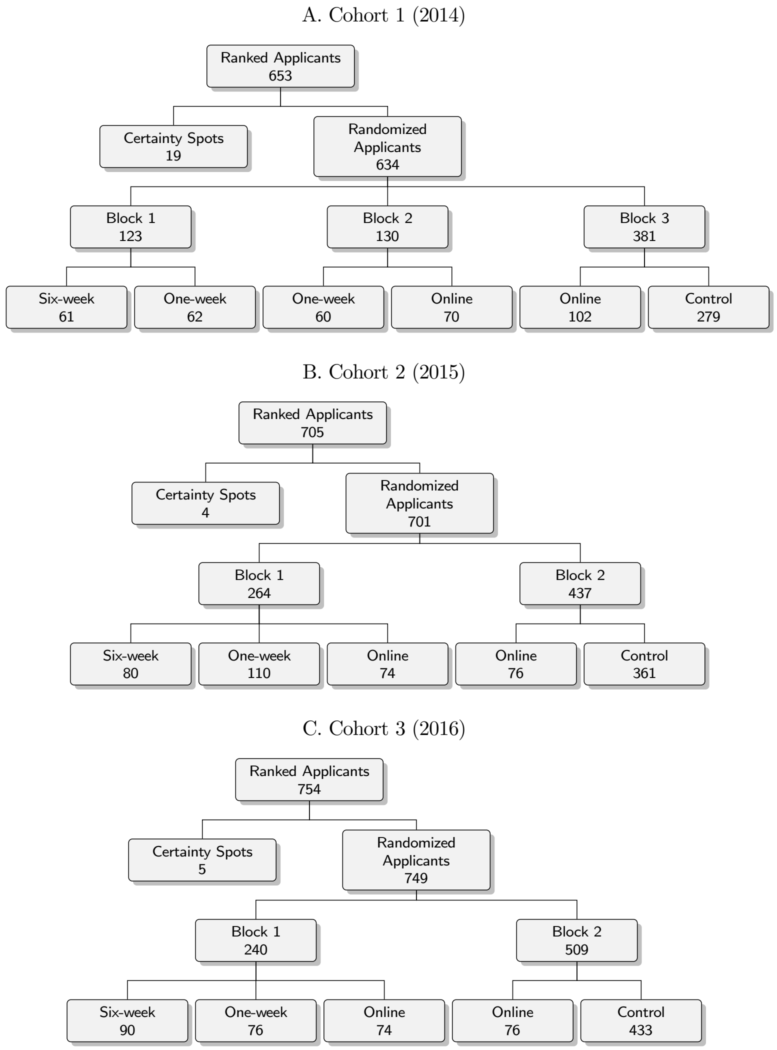

Figure 1

Randomization Design Across Cohorts

Source: HI STEM summer applications from 2014–2016.

Notes: This figure shows the blocked randomization design across the three cohorts (2014–2016). "Certainty spots" are applicants offered admission with certainty and excluded from the experimental analysis. All other ranked applicants were subject to random assignment within block. Block assignment reflects applicant ratings, math scores, and regional priorities, with gender as a stratum within block.

Alt Text (not to be printed): Figure showing blocked randomization design across three cohorts (2014–2016). Each cohort begins with all ranked applicants: Cohort 1 (653), Cohort 2 (705), Cohort 3 (754). A small subset of applicants is allocated to "certainty spots" (19 for Cohort 1, 4 for Cohort 2, 5 for Cohort 3) and is excluded from randomization. The rest are randomized within blocks: Cohort 1 has three blocks (Block 1: 123, Block 2: 130, Block 3: 381); Cohorts 2 and 3 have two blocks each (Block 1: 264/240, Block 2: 437/509). For each block, applicants are randomly assigned to different treatment arms—six-week, one-week, online, or control—with sample sizes per arm detailed for each cohort. The figure structure is consistent with blocked randomization, visually depicted via tree diagrams for each cohort.

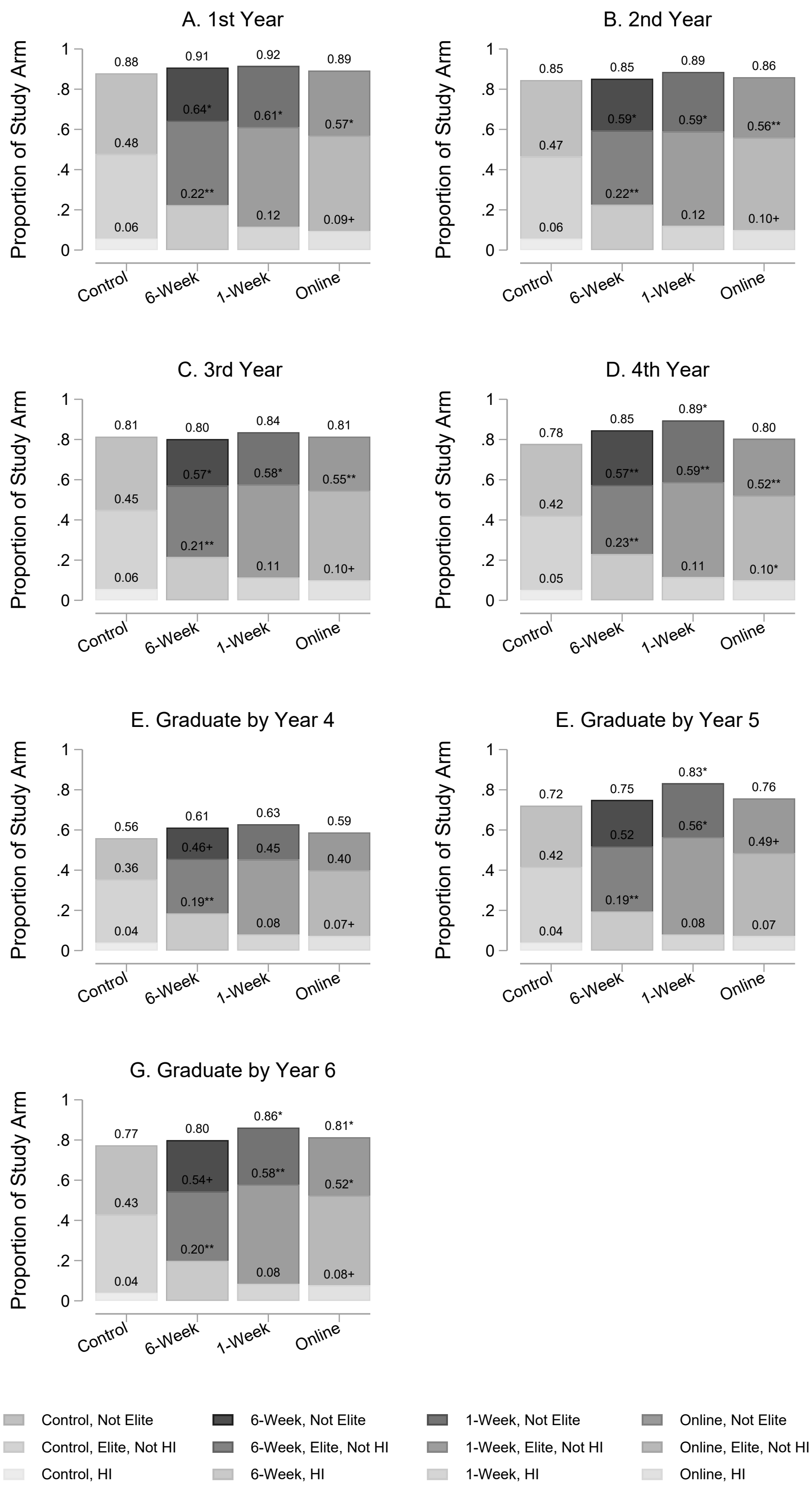
A. 1st Year
B. 2nd Year
C. 3rd Year
D. 4th Year
E. Graduate by Year 4
E. Graduate by Year 5
G. Graduate by Year 6
Proportion of Study Arm
Control
6-Week
1-Week
Online
0.88
0.91
0.92
0.89
0.64*
0.61*
0.57*
0.48
0.22**
0.12
0.06
0.09+
0.85
0.85
0.89
0.86
0.59*
0.59*
0.56**
0.47
0.22**
0.12
0.06
0.10+
0.81
0.80
0.84
0.81
0.57*
0.58*
0.55**
0.45
0.21**
0.11
0.06
0.10+
0.78
0.85
0.89*
0.80
0.57**
0.59**
0.52**
0.42
0.23**
0.11
0.05
0.10*
0.56
0.61
0.63
0.59
0.46+
0.45
0.40
0.36
0.19**
0.08
0.04
0.07+
0.72
0.75
0.83*
0.76
0.52
0.56*
0.49+
0.42
0.19**
0.08
0.04
0.07
0.77
0.80
0.86*
0.81*
0.54+
0.58**
0.52*
0.43
0.20**
0.08
0.04
0.08+
Control, Not Elite
Control, Elite, Not HI
Control, HI
6-Week, Not Elite
6-Week, Elite, Not HI
6-Week, HI
1-Week, Not Elite
1-Week, Elite, Not HI
1-Week, HI
Online, Not Elite
Online, Elite, Not HI
Online, HI

Figure 2

Impact of STEM Summer Program Assignment

Notes: This figure summarizes conditional randomization impact estimates for attendance of and graduation from four-year institutions, split between the HI, other *Barron's* "most competitive" institutions ("Elite"), and remaining four-year institutions from the conditional randomization model. For details on the point estimates and standard errors, see Online Appendix B.

Alt Text (not to be printed): Figure depicting seven groupings (A-G) each containing four vertical stacked bar graphs that show the proportions of participants randomized to the Control, 6-Week, 1-Week, and Online conditions who subsequently attended and graduated from four-year institutions. Each of the four treatment conditions is depicted as a bar. Each bar further is divided into different-shaded segments showing proportion of participants assigned to that condition who attended the HI, other *Barron's* "most competitive" institutions ("Elite, Not HI"), and remaining four-year institutions ("Not Elite"), as well as the cumulative total attending any four-year institution. Groupings A-D track enrollment at institutions over 1-4 years, while Groupings E-G show graduation rates by the end of 4, 5, and 6 years. By Year 6, 77% of participants randomized to the control condition had graduated from a four-year institution; compared with 80% of participants randomized to the 6-Week condition, 86% of participants in the 1-Week condition, and 81% of participants in the Online condition. These proportions can be further broken down into graduation from the HI, from Elite institutions, or from all other institutions, for each of the four treatment models.

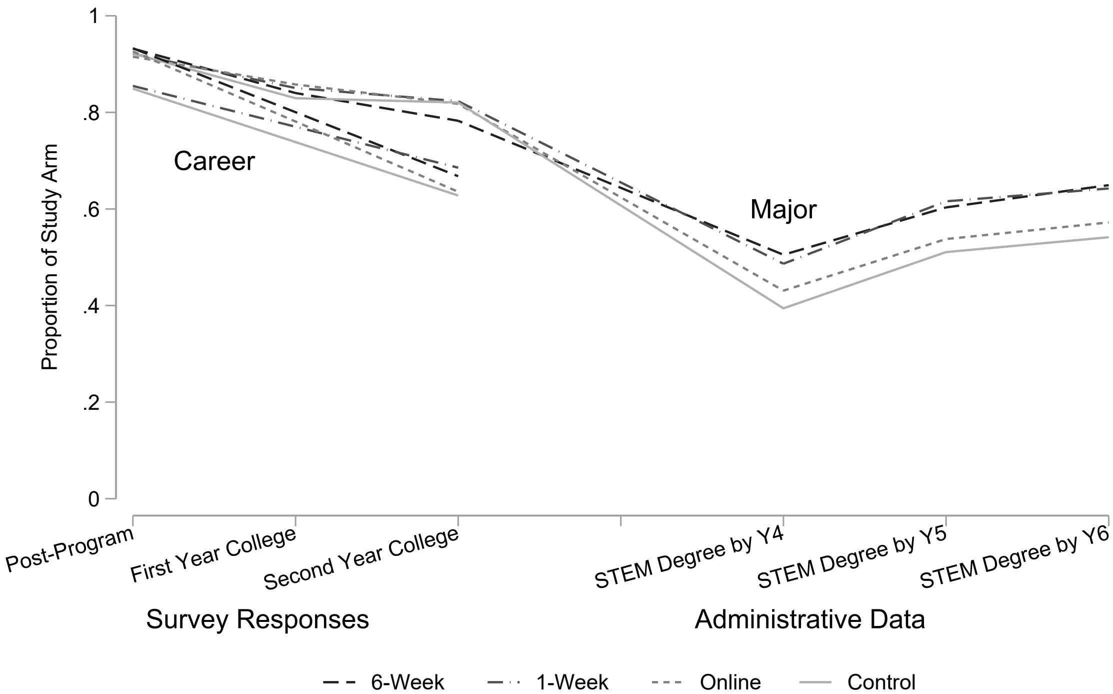

Figure 3

Impact of STEM Summer Programs on Choice of STEM Major and Career Intentions

Notes: This figure shows the proportion of students reporting the intention to major or have a career in a STEM field by program assignment, generated from the conditional random assignment treatment effects. The first set of responses comes from surveys, and the STEM degree information comes from the NSC and HI data. For detailed impact estimates on STEM intentions, see Online Appendix Table B.16.

Alt Text (not to be printed): Line graph showing proportions of students randomized to each of the four treatment programs (Control, Online, 6-Week, and 1-Week) who reported the intention to major or have a career in a STEM field, and how this proportion across different points in time. Points in time include Post-Program, First Year College, and Second Year College (all gathered by survey responses), and also STEM Degree by Year 4, STEM Degree by Year 5, and STEM Degree by Year 6 (gathered from administrative data). All groups follow the same general trajectory of responses – highest interest (between .8-1) reported on surveys post-program, ending with between .5-.6 obtaining a STEM Degree in Years 4-6 – but the control group report the lowest proportion of students expressing intention to have a major or career in STEM for majority of time points.

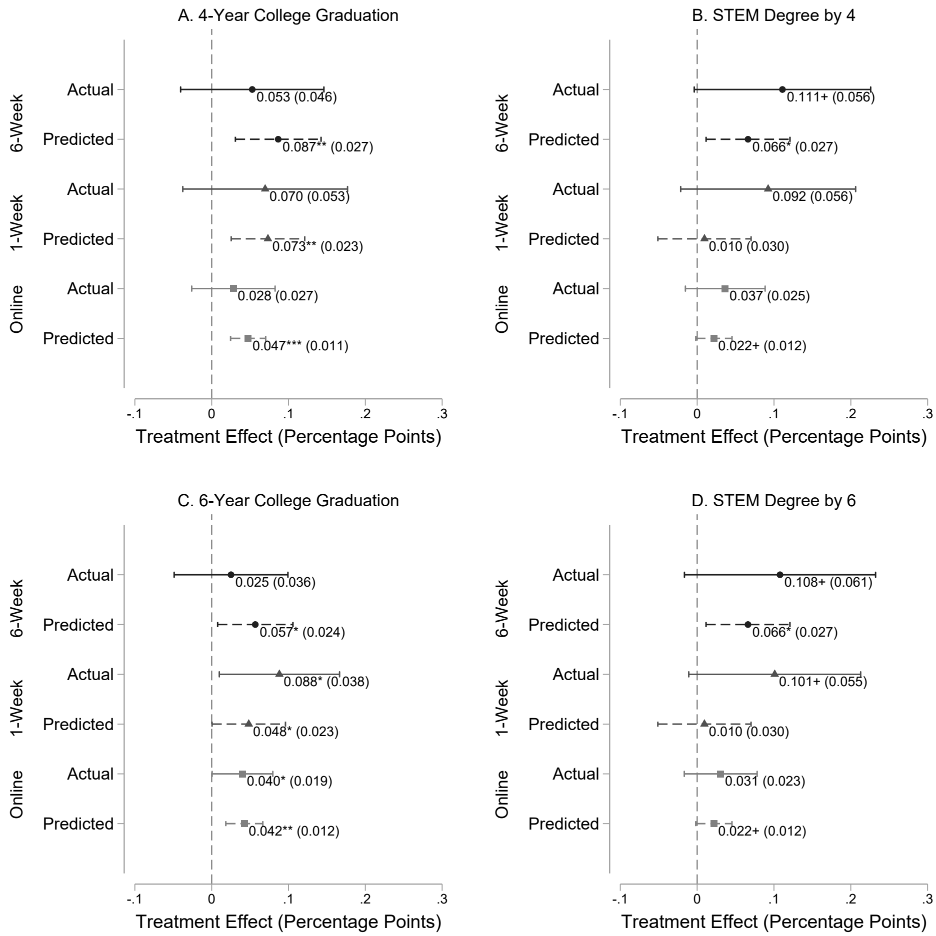

Figure 4

Actual vs. Predicted Graduation and STEM Degree Completion Rates

Notes: This figure compares conditional randomization estimates of the program effects on four-year college graduation within four and six years and STEM degree attainment, marked as "actual," with "predicted" graduation and STEM degree attainment derived on the basis of institution-level characteristics. The institutional-level outcomes are college-level characteristics calculated from IPEDS data in 2013. The "predicted" STEM degree values are the share of STEM degrees among all degrees conferred between July 1, 2012, and June 30, 2013, and do not differentiate between four- and six-year graduation. Values for community colleges and non-college-going respondents are set to zero.

Alt Text (not to be printed): Dot plot displaying treatment effect sizes by condition, on graduation and STEM degree completion rates. The x-axis represents treatment effect in percentage points, centered around zero, with a vertical reference line at zero indicating no effect. The y-axis shows actual and predicted results for each of three treatment conditions (Online, 1-Week, and 6-Week). Each condition is represented by a single dot positioned according to its effect size. Plots are provided to show treatment effects on 4-year College Graduation, STEM Degree by 4 Years, 6-Year College Graduation, and STEM Degree by 6 Years.

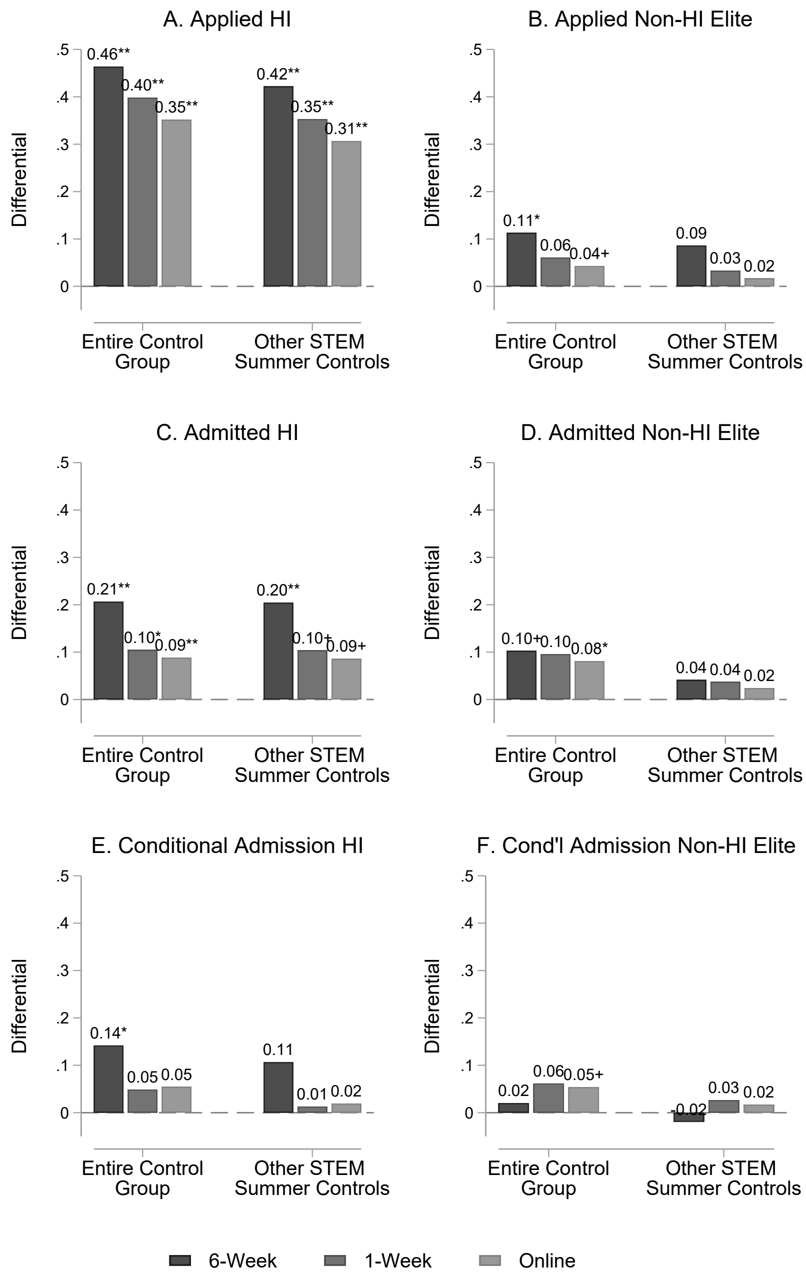

Figure 5

Admissions Outcomes with Alternative Comparison Group

Notes: This figure shows application, admission, and conditional admission to the HI (Panels A–C) and other elite universities (Panels D–F) comparing treated students to all control group students and only to students who report STEM summer experiences. Data on admissions to the HI come from HI administrative records; data on admissions to other universities come from survey responses. For detailed estimates, see Online Appendix Table B.14.41.

Alt Text (not to be printed): Six-panel figure of bar graphs. Panels A,C,E display application, admission, and conditional admission outcomes for HI; Panels B,D,F display the same outcomes for other elite universities. In each panel, bars compare students in the three HI summer programs to two control groups: all control students and a subset of control students who reported having any STEM summer experience but did not attend the HI programs. The y-axis represents the difference in outcomes between each control group and each HI summer program. Bars indicate the estimated magnitude and direction of these differentials. The strongest differentials between program attendees and both control groups are observed when comparing application, admission, and conditional admission to the HI, but almost all panel comparisons show positive differentials between students who attended HI programs and control groups.

Table 1
Baseline Characteristics by Program Assignment

	Full Sample	6-Week	1-Week	Online	Control	Strata-Adjusted Differences: 6-week vs. Control	1-week vs. Control	Online vs. Control
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	0.349	0.403	0.351	0.343	0.339	0.069+	-0.033	-0.019
						(0.041)	(0.035)	(0.026)
Hispanic	0.430	0.407	0.412	0.432	0.440	-0.013	0.008	-0.001
						(0.040)	(0.035)	(0.026)
Native American	0.045	0.056	0.052	0.042	0.041	-0.001	0.005	0.001
						(0.020)	(0.016)	(0.012)
Asian	0.136	0.100	0.140	0.140	0.141	-0.036	0.012	0.014
						(0.027)	(0.025)	(0.019)
White	0.039	0.035	0.045	0.040	0.037	-0.020	0.008	0.002
						(0.017)	(0.015)	(0.011)
Multiethnic	0.358	0.377	0.341	0.354	0.361	0.027	0.002	-0.002
						(0.040)	(0.035)	(0.026)
GPA	3.860	3.911	3.896	3.881	3.830	0.004	-0.003	0.019+
						(0.014)	(0.019)	(0.011)
Free/reduced-price lunch	0.391	0.485	0.455	0.373	0.360	-0.003	0.006	-0.023
						(0.042)	(0.037)	(0.026)
Standardized math score	1.929	2.133	2.135	1.939	1.822	-0.021	-0.030	0.001
						(0.073)	(0.067)	(0.059)
Female	0.404	0.502	0.500	0.502	0.312	0.000	0.000	0.000
						(0.000)	(0.000)	(0.000)
First-generation college	0.242	0.303	0.302	0.225	0.219	-0.007	0.046	-0.025
						(0.037)	(0.033)	(0.023)
p-value						0.830	0.971	0.864
N	2084	231	308	472	1073			

Source: HI STEM summer applications from 2014–2016.
Notes: This table summarizes demographic characteristics, test scores, and GPA for program applicants. Column 1 shows averages taken across the entire sample. Columns 2–5 display means of these traits at baseline by program assignment. Race/ethnicity categories are not exclusive. "First-generation college" is defined as no parental college attendance. Students missing parental college information (N=21) are coded as not first generation. Columns 6–8 report coefficients from regressions of the student characteristic on program offer dummies, including controls for randomization strata (+ p<0.10, * p<0.05, ** p<0.01, ***p<0.001). The *p*-values are from tests of the hypothesis that all coefficients on each program offer variable are zero.

Table 2
Impact of Random Assignment to STEM Summer Programs on Four-Year College Attendance and Graduation

	Attended in Year 1			Graduated by Year 6			
	Any	HI	Elite	Any	HI	Elite	In STEM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Block 1: 6- and 1-week vs. Online							
6-week	0.005	0.107*	0.066	-0.019	0.106**	0.036	0.088
	(0.017)	(0.038)	(0.057)	(0.034)	(0.032)	(0.051)	(0.066)
1-week	0.028	0.038	0.043	0.052	0.024	0.046	0.089
	(0.027)	(0.035)	(0.047)	(0.035)	(0.035)	(0.039)	(0.053)
Online Mean	0.894	0.156	0.633	0.798	0.138	0.573	0.528
Block 2: Online vs. Control							
Online	0.014	0.040+	0.092*	0.040+	0.038+	0.094*	0.030
	(0.017)	(0.020)	(0.037)	(0.019)	(0.019)	(0.039)	(0.022)
Control Mean	0.879	0.057	0.477	0.774	0.040	0.430	0.520
SUR: Treatment vs. Control							
6-week	0.018	0.147**	0.159**	0.021	0.144**	0.130*	0.118+
	(0.041)	(0.047)	(0.061)	(0.054)	(0.044)	(0.063)	(0.063)
1-week	0.042	0.078*	0.136*	0.092*	0.062+	0.139*	0.119*
	(0.036)	(0.040)	(0.056)	(0.046)	(0.037)	(0.057)	(0.058)
Pooled	0.025	0.081**	0.125**	0.054	0.074**	0.120**	0.085+
	(0.029)	(0.028)	(0.043)	(0.036)	(0.026)	(0.044)	(0.044)
Conditional Randomization							
6-Week	0.029	0.167***	0.166*	0.025	0.159***	0.115+	0.115+
	(0.025)	(0.041)	(0.065)	(0.036)	(0.036)	(0.061)	(0.063)
1-Week	0.036	0.059	0.134*	0.088*	0.045	0.148**	0.115*
	(0.030)	(0.038)	(0.059)	(0.038)	(0.037)	(0.054)	(0.055)
Online	0.013	0.038+	0.092*	0.040*	0.036+	0.094*	0.028
	(0.016)	(0.020)	(0.035)	(0.019)	(0.019)	(0.037)	(0.022)
p (6-week = 1-week)	0.730	0.010	0.401	0.056	0.003	0.389	0.996
p (1-week = Online)	0.371	0.528	0.371	0.160	0.789	0.177	0.097
p (6-week = Online)	0.408	0.001	0.191	0.630	0.000	0.676	0.151

Source: NSC records linked to HI STEM summer applications from 2014–2016.
Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including GPA, standardized math score, indicators for race/ethnicity and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014–2016 and passed an initial screening, who were then subject to random assignment. Elite colleges are those ranked as "most competitive" by *Barron's*. Panel A compares students assigned to one of the residential programs to the online program (N = 634). Panel B compares the online program to the control group (N = 1,327). Panel C combines estimates from Panel A and Panel B using SUR. Panel D jointly estimates program effects for all three programs relative to the outcomes of the control group (N = 2,084). The online and control means are adjusted for randomization strata. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001).

Table 3

Impact of Random Assignment to STEM Summer Programs on Potential Earnings

	All Applicants		College Graduates	
	All	By Gender	All	By Gender
	(1)	(2)	(3)	(4)
Block 1: 6- and 1-week vs. Online				
6-week	3,820	6,508	6,639	6,623
	(5,930)	(4,684)	(5,754)	(4,531)
1-week	-596	-3,005	-1,218	-3,213
	(4,113)	(3,115)	(4,254)	(2,985)
Online Mean	114,745	113,468	121,345	115,324
Block 2: Online vs. Control				
Online	7,114*	6,698**	7,371*	6,823***
	(2,606)	(2,047)	(3,267)	(1,601)
Control Mean	99,377	96,755	108,448	100,241
SUR: Treatment vs. Control				
6-week	10,934+	13,213*	14,018*	13,508*
	(6,416)	(5,569)	(7,048)	(6,319)
1-week	6,518	3,701	6,166	3,641
	(5,416)	(4,842)	(5,855)	(5,397)
Pooled	7,865+	7,207*	8,474+	7,195+
	(4,135)	(3,652)	(4,478)	(4,081)
Conditional Randomization				
6-Week	12,467+	14,166**	14,876*	14,923**
	(6,143)	(4,712)	(6,138)	(4,270)
1-Week	5,438	3,079	5,476	3,211
	(4,795)	(3,682)	(5,210)	(3,405)
Online	7,109**	6,609**	7,434*	6,758***
	(2,499)	(1,943)	(3,244)	(1,579)
p (6-week = 1-week)	0.105	0.001	0.030	0.002
p (1-week = Online)	0.687	0.269	0.641	0.259
p (6-week = Online)	0.347	0.090	0.168	0.053

Source: NSC records and College Scorecard data linked to HI STEM summer applications from 2014–2016.

Notes: The notes for this table are the same as in Table 2, but the sample is further restricted to college graduates (Columns 3 and 4), resulting in the following sample sizes: Panel A (N(all applicants) = 634, N(graduates) = 505), Panel B (N(all applicants) = 1327, N(graduates) = 1037), and Panel D (N(all applicants) = 2084, N(graduates) = 1642). Potential earnings come from the College Scorecard and reflect the median earnings by college and major of students who graduated in 2014–2015 and 2015–2016, measured in 2020 and 2021 in 2022 dollars. These measures are linked to study participants via their college and major. If earnings by major are not available because of small sample sizes, earnings by institution are substituted. Students who did not attend college are assigned earnings at the fifth percentile of the college-by-major distribution of earnings in Columns 1 and 2. Columns marked "All" use median earnings, and Columns marked "By Gender" use median earnings for men and women assigned to the matching gender.

Table 4
Impact of Assignment to STEM Summer Programs on Self-Reported College Applications

	Any 4-Year (1)	Number of Apps (2)	Number of Apps Ex. HI (3)	Only One App (4)	Elites Ex HI (5)	Non-competitive (6)	Technical School (7)	Tech School Ex. HI (8)	State Flagship (9)
Block 1: 6- and 1-week vs. Online									
6-week	0.006	-0.785	-0.859	-0.077*	0.074*	-0.007	0.073	-0.028	0.010
	(0.004)	(1.283)	(1.273)	(0.029)	(0.034)	(0.025)	(0.047)	(0.075)	(0.048)
1-week	0.010	-0.407	-0.371	-0.002	0.012	-0.015	0.010	0.029	-0.010
	(0.008)	(0.634)	(0.612)	(0.024)	(0.046)	(0.016)	(0.029)	(0.048)	(0.049)
Online Mean	0.989	9.328	8.666	0.065	0.874	0.040	0.759	0.494	0.448
Block 2: Online vs. Control									
Online	-0.002	0.863+	0.612	-0.034*	0.044+	-0.021+	0.183***	0.081+	0.024
	(0.007)	(0.413)	(0.404)	(0.014)	(0.023)	(0.011)	(0.031)	(0.040)	(0.039)
Control Mean	0.993	8.422	8.024	0.061	0.849	0.061	0.627	0.503	0.505
SUR: Treatment vs. Control									
6-week	0.003	0.083	-0.239	-0.110***	0.118**	-0.028	0.254***	0.053	0.034
	(0.007)	(0.814)	(0.807)	(0.031)	(0.044)	(0.028)	(0.061)	(0.073)	(0.072)
1-week	0.008	0.458	0.242	-0.037	0.057	-0.036	0.193***	0.109+	0.014
	(0.009)	(0.754)	(0.746)	(0.030)	(0.043)	(0.024)	(0.056)	(0.066)	(0.066)
Pooled	0.003	0.507	0.251	-0.055**	0.068*	-0.028	0.204***	0.084	0.023
	(0.007)	(0.567)	(0.560)	(0.021)	(0.033)	(0.019)	(0.044)	(0.051)	(0.051)
Conditional Randomization									
6-Week	0.006	0.148	-0.179	-0.094**	0.113*	-0.035	0.247***	0.039	0.026
	(0.009)	(1.184)	(1.164)	(0.030)	(0.041)	(0.024)	(0.054)	(0.082)	(0.060)
1-Week	0.006	0.342	0.141	-0.041	0.062	-0.035+	0.181***	0.101	0.027
	(0.010)	(0.791)	(0.766)	(0.029)	(0.049)	(0.018)	(0.041)	(0.066)	(0.060)
Online	-0.003	0.822+	0.571	-0.032*	0.043+	-0.022*	0.182***	0.080*	0.026
	(0.007)	(0.406)	(0.397)	(0.013)	(0.023)	(0.010)	(0.029)	(0.038)	(0.036)
p (6-w = 1-w)	0.945	0.794	0.666	0.005	0.043	0.982	0.051	0.121	0.993
p (1-w = Onl)	0.320	0.485	0.517	0.713	0.665	0.387	0.992	0.692	0.994
p (6-w = Onl)	0.191	0.549	0.499	0.030	0.054	0.545	0.163	0.569	0.998

Source: Survey responses linked to HI STEM summer applications from 2014–2016.
Notes: The notes for this table are the same as in Table 2, but the sample is further restricted to survey respondents, resulting in the following sample sizes: Panel A (N = 503), Panel B (N = 809), and Panel D (N = 1418). Data are from surveys conducted at the end of the senior year of high school. The application outcome in Column 1 was were created from a yes/no question for all survey respondents. The number of applications outcomes in Columns 2–4 were calculated for respondents who provided their full list of applications and admissions. The applications to types of institutions outcomes in Columns 5–9 are populated for respondents who provided at least a partial list of applications. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001).

Table 5
Impact of Assignment to STEM Summer Programs on Human Capital

	Plans to Take AP or IB Courses				Soft Skills			
	Any Course	Science	Computer Science	Math	Calculus Question	Life Skills	Study Skills	Confidence
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Block 1: 6- and 1-week vs. Online								
6-week	0.041	0.022	0.025	-0.072	-0.012	0.279*	0.153	0.049
	(0.039)	(0.056)	(0.021)	(0.059)	(0.056)	(0.105)	(0.121)	(0.128)
1-week	0.091*	0.054	0.005	0.040	0.000	0.142	0.110+	0.165
	(0.039)	(0.054)	(0.035)	(0.055)	(0.040)	(0.110)	(0.061)	(0.109)
Online Mean	0.887	0.750	0.129	0.734	0.242	0.048	0.209	-0.107
Block 2: Online vs. Control								
Online	0.026	0.010	0.069	0.028	0.039	0.117	0.273**	-0.007
	(0.015)	(0.041)	(0.049)	(0.044)	(0.044)	(0.098)	(0.065)	(0.110)
Control Mean	0.916	0.736	0.139	0.750	0.220	-0.000	-0.000	-0.000
SUR: Treatment vs. Control								
6-week	0.067	0.031	0.094+	-0.044	0.027	0.395**	0.425***	0.041
	(0.043)	(0.067)	(0.056)	(0.068)	(0.066)	(0.126)	(0.123)	(0.156)
1-week	0.117**	0.064	0.074	0.068	0.039	0.258*	0.381***	0.157
	(0.039)	(0.065)	(0.056)	(0.066)	(0.067)	(0.115)	(0.110)	(0.158)
Pooled	0.073*	0.037	0.080+	0.017	0.035	0.246**	0.354***	0.068
	(0.032)	(0.053)	(0.046)	(0.053)	(0.053)	(0.093)	(0.090)	(0.126)
Conditional Randomization								
6-Week	0.063	0.026	0.097+	-0.050	0.029	0.378*	0.441**	0.028
	(0.043)	(0.071)	(0.050)	(0.071)	(0.066)	(0.141)	(0.125)	(0.165)
1-Week	0.112*	0.058	0.078	0.060	0.036	0.267+	0.369***	0.134
	(0.041)	(0.067)	(0.057)	(0.067)	(0.051)	(0.143)	(0.088)	(0.143)
Online	0.027+	0.008	0.073	0.025	0.039	0.116	0.268***	-0.017
	(0.015)	(0.042)	(0.047)	(0.042)	(0.038)	(0.100)	(0.058)	(0.105)
p (6-w = 1-w)	0.016	0.496	0.627	0.062	0.894	0.159	0.343	0.359
p (1-w = Onl)	0.036	0.349	0.876	0.502	0.937	0.161	0.135	0.154
p (6-w = Onl)	0.380	0.755	0.247	0.205	0.860	0.016	0.132	0.732

Source: Survey responses linked to HI STEM summer applications from 2014–2016.
Notes: The notes for this table are the same as in Table 2, but the sample is further restricted to survey respondents, resulting in the following sample sizes: Panel A (N = 570), Panel B (N = 905), and Panel D (N = 1576). Data are from surveys conducted at the end of the program summer. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 *** p<0.001).

[i] The STEM wage premium likely reflects selection into STEM fields by individuals with high earning potential but remains even when student background is accounted for (Arcidiacono 2004; Altonji, Blom, and Meghir 2012; Hastings, Neilson, and Zimmerman 2013; Kinsler and Pavan 2015; Altonji, Arcidiacono, and Maurel 2016; Kirkeboen, Leuven, and Mogstad 2016; Lovenheim and Smith 2023).

[ii] In 2011, there were over 250 federal programs and $3.4 billion invested in cultivating the STEM pipeline (Granovskiy 2018), including over $1 billion in funding through the National Science Foundation (NSF) with the specific goal of increasing diversity and representation. Approximately three-quarters of this federal investment supports undergraduate and graduate education and training, with K–12 initiatives receiving less support (authors' calculations from Granovskiy 2018).

[iii] Some summer programs to increase representation in STEM-adjacent fields, including the American Economic Association Summer Program (Price 2005; Becker, Rouse, and Chen 2016) and the Robert Wood Johnson Foundation Summer Medical and Dental Education Program (Cosentino et al. 2015), have been rigorously analyzed and found to increase representation in their focus areas. However, although these evaluations offer more rigorous comparisons to nonprogram students than many other studies, they are not randomized trials, and they focus on STEM-adjacent fields (e.g., economics, health professions). These works generally find that program participation leads to greater success in the focus field.

[iv] Technical institutions are universities that focus on STEM fields and often include in their name the words "Polytechnic," "Institute of Technology," "Tech," or "Technological." Some examples outside the Northeast include Harvey Mudd College, Georgia Institute of Technology, and Rose-Hulman Institute of Technology.

[v] This program is no longer operating.

[vi] STEM majors are computer and information science, engineering, engineering technologies, biological and biomedical sciences, mathematics and statistics, physical sciences, and science technologies. Due to lack of reporting to the NSC, information is missing for 12–15 percent of the degree recipients, depending on the time horizon. To address this data missingness issue, we assume that bachelors of science represent STEM majors for those without information on majors. We otherwise categorize majors as non-STEM. We also explicitly display results by the rate of missingness of degree information and show upper and lower bounds on the results derived when we categorize all missing degrees as either non-STEM or STEM. The HI always reports the degree field. See Table B.9 for details.

[vii] For more details on covariate balance, see Online Appendix Tables A.4–A.6, which show that within strata, there are few differences by background characteristics across treatment and control conditions *within* randomization blocks.

[viii] Applicants to the STEM summer programs submit standardized test scores from various exams, the most common being the PSAT. We use information on national test score means and standard deviations to convert each test score to a z-score, which allows us to combine across several standardized exams including the PSAT, SAT, ACT, and PLAN. To compare the program applicants' academic backgrounds to those of students typically admitted to the HI, we standardize by the HI's 25th percentile math SAT of the incoming freshman cohort in the same period as the STEM summer program experiment, which was 740, or 1.91 in standard deviation units. Thus, the average scores of program applicants were in line with the scores of incoming students at the HI.

[ix] In practice, the test score floor was binding only in the 2015 random assignment. In 2014, no test score floor was imposed, and blocks were determined solely by rating. In 2016, all applicants who submitted test scores scored above the floor. In 2015, a small number of applicants who had high rating scores but did not pass the test score floor were moved to Block 2 and randomized within that block. Applicants with missing test scores (in all years) could be placed in Block 1. In all cases, block assignment occurred before randomization.

[x] The regional preferences are geographic preference indicators for regions of the country that are preferred, neutral, or downweighted. These strata do not include block because all estimation occurs within blocks in this strategy.

[xi] It is not possible to simply estimate a combined treatment effect vs. the control group via the conditional randomization estimates we describe below. This is because the conditional randomization estimates rely on the blocking structure and the presence of the online treatment in both blocks to make the comparison to control.

Estimates that pool the treatments and leave the blocking structure intact will essentially replicate the online-to-control comparison because there is no area of common support to compare the residential programs to the control group. An alternative would be to pool treatments and remove the blocks from the estimate of the combined treatment effect. However, this estimate would discard essential information from the blocks and would not return a meaningful estimate of program effects.

[xii] These randomization strata are different from the S_{ij} used above as R_{ik} include randomization block whereas S_{ij} do not.

[xiii] For simplicity, our expectation notation omits conditioning on covariates other than randomization strata. The covariates increase precision but are not fundamental to the randomization schema.

[xiv] Technically, it is a conditional average treatment effect because it holds only for those who applied to the programs, not for a representative population.

[xv] Note that the college graduation experiences of these students may have been affected by the COVID-19 pandemic; however, this would have influenced both the treatment and control groups. See Online Appendix Figure A.1 for a timeline of each cohort's expected college enrollment.

[xvi] If a student is missing information on major but her degree is designated a bachelor of science (or some variation thereof, such as BS) rather than a bachelor of arts or BA, then we count it as a STEM degree.

[xvii] Online Appendix Tables B.13 and B.12 report the same results estimated with missing majors counted as non-STEM and STEM majors, respectively. The coding of missing degrees does not affect the overall finding that the programs increased STEM degree attainment, though the various coding schemes do change the magnitudes and statistical significance of some coefficients. However, the pattern of the changes is not systematically smaller or larger.

[xviii] Summary earnings measures by major and institution surely conflate correlations between selection into institutions and causal effects of colleges or majors. However, even in the case when research shows little return to college quality (Dale and Krueger 2002), there tend to be returns for URM (Dale and Krueger 2014). This gives credence to the idea that the students in our sample could achieve the potential earnings from the Scorecard.

[xix] For students who attended community colleges or did not attend college, we substitute zeroes instead.

[xx] These could come via listing program participation, a recommendation letter from an HI-affiliated individual, or inclusion of the program's written evaluation of the student in the college application.

[xxi] These include similar programs at other colleges and universities, nonprofit programs such as Girls Who Code, internships in STEM or medical fields, and STEM summer courses.

[xxii] See Online Appendix E for details on the variables included and how we generated indices from individual variables. Online Appendix E also includes additional survey responses not discussed in the text.