

NBER WORKING PAPER SERIES

STEM SUMMER PROGRAMS FOR UNDERREPRESENTED
YOUTH INCREASE STEM DEGREES

Sarah R. Cohodes
Helen Ho
Silvia C. Robles

Working Paper 30227
<http://www.nber.org/papers/w30227>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2022

We are grateful to staff of and applicants and participants in the summer programs that are the subject of this paper, as well as the host institution and its institutional research office. Because the institution is anonymous we unfortunately cannot individually thank the many people who supported this project by name. Their commitment was instrumental to this experiment, as was feedback and suggestions from Joshua Angrist, Lawrence Katz, and Amanda Pallais. Rebecca Johnson, Stephanie Owen, and seminar and conference participants at APPAM, AEFPP, Colby College, the Opportunity and Inclusive Growth Initiative at the Minneapolis Federal Reserve Bank, and the University of Michigan 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 and the identifying number is: AEARCTR-0002888. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Sarah R. Cohodes, Helen Ho, and Silvia C. Robles. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

STEM Summer Programs for Underrepresented Youth Increase STEM Degrees
Sarah R. Cohodes, Helen Ho, and Silvia C. Robles
NBER Working Paper No. 30227
July 2022
JEL No. I21,I24,J15,J16

ABSTRACT

The federal government and many individual organizations have invested in programs to support diversity in the STEM pipeline, including STEM summer programs for high school students, 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 at an elite, technical institution. The STEM summer programs differ in their length (one week, six weeks, or six months) and modality (on-site or online). Students offered seats in the STEM summer programs are more likely to enroll in, persist through, and graduate from college, with gains in institutional quality coming from both the host institution and other elite universities. The programs also increase the likelihood that students graduate with a degree in a STEM field, with the most intensive program increasing four-year graduation with a STEM degree attainment by 33 percent. The shift to STEM degrees increases potential earnings by 2 to 6 percent. Program-induced gains in college quality fully account for the gains in graduation, but gains in STEM degree attainment are larger than predicted based on institutional differences.

Sarah R. Cohodes
Teachers College
Columbia University
525 West 120th Street
New York, NY 10027
and NBER
cohodes@tc.columbia.edu

Silvia C. Robles
Mathematica
220 E. Huron St.
Ann Arbor, MI 48104
SRobles@mathematica-mpr.com

Helen Ho
The People Lab
Harvard Kennedy School
79 John F. Kennedy Street
Cambridge, MA 02138
helenho@hks.harvard.edu

A data appendix is available at <https://www.nber.org/data-appendix/w30227>

1 Introduction

Black and Hispanic workers are underrepresented in the high-wage, college degree-holding STEM—science, technology, engineering, and mathematics—workforce (National Science Board, 2021). This under-representation contributes to racial and ethnic wage gaps (Altonji et al., 2016).¹ Racial inequality and the lack of diversity in STEM fields are also detrimental to the quality of innovation (Parrotta et al., 2014; Hofstra et al., 2020), and overall economic growth and the presence of racial economic inequality in the United States (Cook et al., 2021; Hsieh et al., 2019).

Under-representation in the STEM workforce is preceded by disparities in who attains a college degree in STEM. About 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 of the college-age population in the United States, respectively (National Science Board, 2022). The disparity in STEM degrees is not due to differences in interest. Upon entering college, underrepresented minority (URM) students plan to major in STEM fields at similar rates to their White peers, but they are more likely to switch away from a STEM field or leave college (Rieggle-Crumb et al., 2019).

Given that STEM preparation and college access are shaped prior to college entrance, STEM-focused enrichment programs for high school students are promising vehicles to reduce disparities in STEM degree attainment. However, we know little about the efficacy of such programs. The existing evidence primarily relies on survey assessments and on observational studies whose findings can be substantially driven by selection bias (unobserved differences between participants and non-participants) and often focus only on short-term outcomes (Kitchen et al., 2018a; Kitchen et al., 2018b; Bradford et al., 2021).² An exception is Robles’ (2018) prior investigation of one of the three programs we examine here (the six-week program). Using data on earlier cohorts and a selection-

¹The STEM wage premium likely reflects selection into STEM fields by individuals with high earning potential, but STEM earnings premiums remain even when controlling for student backgrounds (Altonji et al., 2012, 2016) or estimating returns within a discrete choice model (Arcidiacono, 2004; Kinsler and Pavan, 2015).

²Some summer programs to increase representation in STEM-adjacent fields have been rigorously analyzed and found to increase representation in their focus areas, including the American Economic Association Summer Program (Price, 2005; Becker et al., 2016) and the Robert Wood Johnson Foundation Summer Medical and Dental Education Program (Cosentino et al., 2015). However, although they are more rigorous about comparisons to non-program students than many other studies, these evaluations are not randomized and focus on STEM-adjacent fields (e.g., economics, health professions). They generally find that program participation leads to greater success in the focus field.

on-observables design, she found that access to a STEM summer program increases matriculation at the host institution, graduation rates, and likelihood of graduation with a STEM degree. However, although this study has the benefit of a long time horizon and detailed administrative data, it still cannot fully account for selection into the program. Thus, to better understand the impact of STEM programs for high school students, we conducted a randomized controlled trial of a suite of summer programs targeted at enhancing the pipeline of underrepresented students in STEM degrees and careers, following students from their application to the programs through college graduation with (or without) a STEM degree.

This study provides the first evidence from a randomized controlled trial on the impact of a STEM-focused summer program on college matriculation, completion, and graduation with a STEM degree. Three cohorts of high-achieving, STEM-interested students 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. The programs were held at the Host Institution (HI), an elite technical university in the Northeast. They differed in their modality and intensity: six weeks full-time on-site, one week full-time on-site, or six months with periodic meetings online. Students were selected into the randomization pool based on their academic preparation as well as 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. 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 three programs increase the likelihood of application to, acceptance at, and enrollment in the HI, with the largest magnitudes coming from the six-week program. The STEM summer programs also increase enrollment at Barron’s most competitive colleges: for the six-week on-site program, this operates entirely through the HI; for the one-week and online programs, enrollment

in the most competitive institutions is split between the HI and other elite schools, though some of the effects are not statistically significant. The STEM summer programs also induce students to persist in college. Almost all students in the control group (87 percent) attend a four-year college immediately after high school graduation, and the programs increase this a small amount (2 to 4 percentage points, and not statistically significant). However, by the fourth year of college, enrollment among the control group drops to 75 percent, most likely reflecting students dropping out of college or taking time off before completion. In contrast, those offered a seat in any one of the three STEM summer programs are 3 to 12 percentage points more likely to still be enrolled in a four-year college, largely driven by enrollment in Barron's top ranked colleges.

The STEM summer programs also increase on-time college graduation. Only 53 percent of students in the control group graduate within four years from any four-year school, despite being an academically talented group. The STEM programs increase this by 8 percentage points (six- and one-week on-site programs) and 1.6 percentage points (online program) though the differences are not statistically significant. Again, most of the gains for the six-week program operate through increased graduation from the HI; for the one-week and online programs, college graduation increases are shared by the HI and other highly ranked institutions. Graduation impacts are larger and statistically significant with a five-year window for graduation, but this may be due to sample composition as graduation information is only available for two of three cohorts due to a shorter time horizon.³

Degree gains are entirely in STEM fields, reflecting both an overall increase in the number of degrees and a shift to STEM fields among graduates. In the control group, 34 percent of students graduate within four years with a STEM degree—64 percent of degree recipients. The six-week program increases the rate at which students graduate with a STEM degree to 50.7 percent, 46.8 percent for the one-week program, and 37.2 percent for the online program (the latter is not significant). Much of the shift to STEM occurs with degrees at the HI, but both the six-week and one-week programs increase receipt of STEM degrees at non-HI institutions as well when looking at five-year graduation, though these differences are not statistically significant. The gains for

³Table 3 presents positive six-year graduation effects as well, but they reflect only a single cohort, so the discussion does not emphasize them. Online Appendix Table B.2 shows how graduation effects fluctuate across cohorts.

the online program are split between an increase in STEM degrees at the HI and an increase in non-STEM degrees at other institutions. The shift in composition of majors toward STEM induces potential earnings increases of 2 to 6 percent, which is likely an underestimate because it only accounts for changes in majors, not an increase in graduation.

Evidence indicates the programs' effect on degree completion is due to the shift in institutional quality that they induce. The increases in overall graduation are the same as what would be predicted by the shifts in institutional quality, as measured by institution-level graduation rates. The programs potentially achieve upgrades in institutional quality by improving information students have about colleges and the college application process. We use survey data to explore other mechanisms, such as improvements in study and independent living skills and high school preparation.

This paper makes three main contributions. First, we add to the evidence on STEM degree attainment as well as 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 et al. (2010); Bettinger (2010); Price (2010); Fairlie et al. (2014); Fischer (2017); Griffith and Main (2019), student beliefs in their own capability and signals from grades (Astorne-Figari and Speer, 2019; Kaganovich et al., 2021; Owen, 2020; Kugler et al., 2021; Owen, 2021) and institutional effects (Griffith, 2010; Arcidiacono et al., 2016). Less attention has been paid to the preparatory experiences that may shape college attendance and major choices, despite the potential influence of pre-college 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. In the United States, Darolia et al. (2020) found that exposure to more STEM courses in high school does not increase STEM degree attainment in college, while De Philippis (2021) found that in the United Kingdom, such exposure increases the likelihood of male students majoring in STEM, and in Denmark, Joensen and Nielsen (2016) found an increase only for female students. Although the differences in these findings may be due to differences in context, it is also possible that broad programs that do not

specifically focus on underrepresented students or 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 institutional quality, and the potential for college education to reduce racial economic inequality in the United States. There are large gaps in college enrollment by family income in the United States. (Bailey and Dynarski, 2011; Chetty et al., 2020; Dynarski et al., 2021). And, in addition to whether students enroll in college, there are differences in the type and quality of institution they enroll in (Baker et al., 2018; Gerber and Cheung, 2008). College enrollment and selectivity trail even for high-achieving, underserved students (Hoxby and Avery, 2013; Dillon and Smith, 2017) resulting in “undermatch,” that is, when students who could succeed at selective institutions do not apply (and thus cannot enroll). The college a person attends can influence the likelihood that the student graduates (on time), the likelihood the student graduates with certain degrees, and the student’s future employment and earnings (Hoekstra, 2009; Cohodes and Goodman, 2014; Zimmerman, 2014; Goodman et al., 2017; Chetty et al., 2020; Bleemer, 2021). Interventions prior to college application can influence enrollment and the specific institution students enroll in (Avery, 2010, 2013; Carrell and Sacerdote, 2017; Castleman and Goodman, 2018; Andrews et al., 2020; Dynarski et al., 2021). Similar to effective college counseling and informational interventions that modify college application and enrollment behavior, the STEM summer programs we examine happen at a crucial time: students are seriously considering college but have not yet applied. However, the STEM summer programs we focus on differ in their intensity and focus on STEM.

Finally, this paper is relevant to a large literature on the impacts of affirmative action (or lack thereof) in college admissions. The STEM summer programs do not introduce group-based preferences in college admissions—policies that are typically the focus of the affirmative action literature in economics. They do, however, focus on populations exhibiting a broad definition of need that includes identifying with historically underrepresented groups, and aim to increase access to STEM fields and elite universities. However, much of the literature on affirmative action is concerned with “mismatch”—the idea that 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 et al. (2016) found some evidence of mismatch, Bleemer (2022) found college and earnings benefits for URM students induced to attend more selective University of California campuses due to affirmative action. Our work adds to the evidence that when URM students are induced to attend high-quality institutions, they reap the benefits of those institutions and are successful, in contrast to the predictions of mismatch theory, though we note that the relevant sample here has significant academic preparation for college.

The paper proceeds as follows. Section 2 describes program background and context, including more details on the interventions; Section 3 details the data; and Section 4 explains the study design and estimation methods. Results are reported in Section 5, with a discussion of potential mechanisms in Section 6. Section 7 concludes.

2 STEM summer programs at the Host Institution

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.” Programming includes outreach to the local community with initiatives designed for elementary and secondary students, as well as 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, STEM curriculum, tours of a college campus and research facilities, and college admissions information. Recruitment is national. All programs cover student costs except for transportation to and from the HI. The programs are funded by the HI, with some funding due to earmarked charitable gifts. High-achieving students in any geographic region can be recruited, as long as they are U.S. citizens or permanent residents. One source of student information used in direct mailings for recruitment is the PSAT, though test scores are not a prerequisite for admission.

We describe each summer program below as it existed in the summers of 2014-2016, the period over which randomization occurred. All of the outreach office’s programs offer similar experiences

that are 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 work spaces at the HI; attend workshops with leaders of industry and 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. About 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 one 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 to 120 students participated in this program each year⁴.
3. *Online program:* The online treatment draws on communications technology to serve students. The six-month program provides a platform for multimedia interaction between students and instructors, staff at the HI, 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. Finally, students spend five days on campus presenting their

⁴This program is no longer operating.

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. About 150 to 175 students participate in the online program. The summers 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 are 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 an applicant has no exposure to STEM focused programming. Many of students in this group participate in alternative summer programming, both STEM-focused and otherwise, such as programs administered by other universities or organizations like Girls Who Code or Leadership Enterprise for a Diverse America. However, many also work or study over the summer in lieu of specialized programming.

3 Data and descriptive statistics

This study measures the effectiveness of STEM summer programs for high-achieving, underrepresented high school students using data from a randomized experiment of admission to these programs for 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, records of college attendance and graduation, and survey data, and also describe the characteristics of program applicants.

3.1 Data

Data for our key analyses come from two main sources: program application and admissions information from the HI and college attendance and enrollment information from the National

Student Clearinghouse (NSC). All applicants in the three cohorts from 2014 to 2016 were admitted via conditional random assignment, and 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. Background information on the randomized sample comes from program applications which include demographic information, academic qualifications, and essays, as well as a baseline survey. The outreach office provided details on offers of admission and which students ultimately participated in the programs, as well as details on ratings of applicants' files and details about the admissions process.

Outcome data come from records of college enrollment provided by the HI institutional research office and the NSC; the former also provides information on applications and majors at the HI. Almost all of the applicant pool appears in the HI or NSC college enrollment data, and since 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.13). College outcomes include graduation as of spring 2021, included in both the HI and NSC data, which means we have the potential to observe four-year college graduation for all cohorts, five-year college graduation for the first two cohorts, and six-year graduation for the first cohort. The NSC data also included information on students' majors, which we categorize as either a STEM field or not.⁵ Figure 1 shows data availability and progress through college for each of the three cohorts, assuming on-time progression.

3.2 Surveys

The study also collected periodic survey data from the study sample. Longer surveys were conducted in the fall shortly after the program summer, May of students' senior year of high school, and in the spring of students' sophomore year of college. Shorter, more frequent surveys kept track of college enrollment and students' ultimate or intended college major. Respondents received Amazon

⁵Due to lack of reporting to the NSC, information is missing for 12 to 15 percent of degree recipients depending on the time horizon. To address this, we assume that bachelor of science degrees represent STEM majors for those missing information on majors. We also explicitly display results for rates of missing degree information, as well as show upper and lower bounds from categorizing all missing degrees as non-STEM or STEM, respectively. The HI always reports degree field. See Table 4 for details.

gift cards if they participated (not contingent on answering all questions or on treatment assignment), with larger incentives for participating in the longer surveys (\$25) and smaller incentives for participating in the shorter surveys (\$5).

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 information 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 outcome measures from the surveys by “family” of outcomes 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.13 shows response rates for each survey. For example, for the first long follow-up survey in the fall after program participation, treatment groups’ response rates ranged from 85 to 90 percent; 65 percent of the control group responded to this survey. Differences in response rates are not surprising, given that treatment students were more likely to have positive associations with the program office or be enrolled in the HI, which was 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 C.1). Impacts on attendance and graduation are generally quite similar when restricted to the sample of survey respondents, perhaps a little larger for survey respondents. Adding inverse propensity weights based on demographic characteristics and program assignment has an inconsistent impact on the treatment effects. In some cases, it makes the estimates of college impacts for survey responders look more similar to those of the full sample than those of the unweighted respondents; in others, it makes the estimates for survey respondents look more divergent from those from the full sample. Thus, it is not clear that propensity to respond predicts program effects in a meaningful way.

Following Dutz et al. (2021), we use unweighted survey responses but caution that the sample is not fully representative; thus, we consider findings based on the surveys to be suggestive.

3.3 Descriptive statistics

Table 1 reports demographic and background academic information for the randomized sample. As described below, randomization included design strata, so we do not expect treatment and control groups to be exactly similar on all characteristics, though we show in Columns 6 through 8 that the treatment group has fewer strata-adjusted differences than the control group.⁶ Overall, almost all randomized applicants identify as a member of a group underrepresented in higher education and STEM fields, with 35 percent of the sample identifying as Black, 43 percent identifying as Hispanic, and 4 percent identifying as Native American (note that these categories are overlapping as students reported more than one race or ethnicity). About one-quarter of the group are first-generation college students, which we define as having no parents who ever attended a four-year college. Program applicants have strong academic backgrounds. The average grade point average (GPA) is 3.86, and the average standardized math test score is two standard deviations above the national mean.⁷ The largest contrast between treatment and control groups is by 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 include gender.

4 Research design

STEM summer program applicants were randomized to receive an offer to participate in one of the three summer programs or were randomly assigned to a control group. This section describes the process through which applicants were randomized and how the analysis estimates program effects

⁶For more details on covariate balance, see Online Appendix Tables A.14 through A.16, which show that within strata, there are few differences by background characteristics across treatment and control conditions *within* randomization blocks.

⁷Applicants 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 to the HI, we standardized the HI's 25th percentile of the math SAT in the same time period as the STEM summer program experiment, which was 740, or 1.91 in standard deviation units. Thus average applicant scores of program applicants were in line with scores of incoming students at the HI.

based on that randomization process.

4.1 Selection

Selection into the programs is a multi-step process, described in more detail in Online Appendix A. Each year, after an initial screening, the outreach office sent approximately 600 to 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 for academic preparation and personal circumstances, based on 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 and engineering backgrounds
3. The applicant's high school historically sent less than 50 percent of its graduates to four-year colleges
4. The applicant attended a high school that presented challenges for success at an elite, urban university (e.g., rural high school or a high school with a predominantly URM population)
5. The applicant was a member of a group that is underrepresented in the study and fields of science and engineering (African American or Black; Hispanic or Latino/a; or Native American)

Additionally, the outreach office 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 admission to these applicants to support national representation; we call these cases “certainty spots,” and they are excluded 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 applicant ratings and regional priorities; the rankings were used to allocate students to random assignment blocks as described in Section 4.2 below. Notably, because randomization occurs in a group designated as top applicants from a pool of more than 2,000 applicants, program effects estimated here may not apply to all applicants or to others outside the selective applicant pool. The program, if offered to any high school senior, might have very different (or no) impacts on those less academically prepared.

4.2 Randomization

The selection process above created a pool of 600 to 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 based on the rating variable that took into account application ratings and regional priorities. An overview of the randomization scheme is in Figures 2a-2c. Details on the randomization process for each cohort are 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. 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 based on operational considerations. This randomization scheme formed the research design in cohorts 2 and 3 of the experiment. Cohort 1, in 2014, underwent a slightly different design, where students were differentiated to a greater extent and randomly assigned within three blocks. Results are very similar including and excluding the first cohort (see Online Appendix Table A.6). We include the 2014 cohort to increase statistical precision, despite the slightly different underlying randomization structure, which we account for with randomization

strata.

The crucial component of our 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 we assume that if we control for randomization strata (based on application year, gender, block, and regional preferences), we fully account for differences across the blocks and can compare 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 block with known variables. We test this assumption in multiple ways, including controlling for the rating variable directly, which we describe below after presenting our main estimation strategy.

4.3 Estimation

We use random assignment to program offers to estimate the causal effect of assignment to one of the three STEM summer programs, as follows:

$$Y_i = \beta_1 6week_i + \beta_2 1week_i + \beta_3 Online_i + \sum_{j=1}^J \delta_j R_{ij} + X_i' \gamma + \epsilon_i \quad (1)$$

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 indicators for random assignment to an offer of treatment for each of the three programs. The parameters of interest are the β coefficients, which reflect the intent-to-treat estimate of assignment to one of the three programs. Most students attend their assigned program if offered a spot, with 87 percent of students accepting a seat at the six-week program, 85 percent at the one-week program, and 77 percent at the online program (Online Appendix Table A.2). Very few students end up participating in a different program, and the outreach office did not offer any spots to students in the control group.⁸ A vector of student-level control variables, X_i , including GPA, standardized math scores, race/ethnicity, and free and reduced-price lunch status, increases precision. Gender is accounted for in the randomization strata described below. We use

⁸Assignment closely parallels participation, and our estimates will be quite close to treated-on-the-treated impact estimates. Nevertheless, we also instrument program attendance with program assignment and present treatment-on-the-treated estimates, as well (Online Appendix Tables B.3 and B.4).

heteroskedasticity robust standard errors.

Key to our estimation strategy is the inclusion of a set of control variables, or risk sets, R_{ij} , which are indicators representing randomization strata. Randomization strata represent gender, regional preferences,⁹ and randomization block, and are formed within randomization year. Offers are randomized within these strata. Students are assigned to randomization blocks based on 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 make sure that participants are broadly representative of the United States as a whole). Our estimation method compares students within the same cohort, gender, regional preference, and randomization block. We show in Table 1, Columns 6 through 8, that once we control for randomization strata there are few differences in student characteristics. Online Appendix Tables A.14 through A.16 show that demographic characteristics are balanced within randomization strata.

The fundamental assumption behind our randomization strategy is that once we control for randomization block, we are controlling 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. Including the online program in both randomization blocks provides the link that makes it possible to estimate this comparison. Because assignment to block is based completely on known, observable variables, controlling for randomization block should control for all differences between the two groups.

Nevertheless, because this differs from complete random assignment, several alternative estimation strategies support the interpretation that the estimates represent causal effects, detailed in Section 5.5. Specifically, we verify this strategy is successful by showing our results controlling for the rating variable in lieu of the block-based randomization strata, which shows that the blocking strategy fully accounts for the factors that determine assignment to randomization group, removing selection bias. Because the blocks are constructed out of known parameters and the analysis fully accounts for them, the modified random assignment structure comes close to estimates under

⁹These are geographic preference indicators for regions of the country that are preferred, neutral, or down-weighted.

complete random assignment.

There is another potential threat to the validity of the modified randomization scheme: If there are differential returns to the program for different types of applicants—that is, if there are heterogeneous treatment effects—our results may not be fully representative because the control group only includes relatively lower-ranked students. Online Appendix A.1.6 explores this concern and shows little evidence of heterogeneous treatment effects.

Section 5.5 presents several additional robustness checks. Because randomization possibilities are constrained to only those within the strata discussed above (i.e., complete randomization is not possible), we also present our estimates using 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). 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 a 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.

5 Results

This section details how the programs increase both college enrollment and college graduation and shift the type of institution that students attend to higher-quality colleges. The programs increase enrollment in and graduation from the HI as they were designed to do. However, it could be the case that the programs are simply shifting enrollment and graduation from similar institutions to the HI. We explore the impact of a STEM summer program offer on enrollment and graduation from four-year colleges more broadly, as well as by college quality and type. Because the NSC, which tracks enrollment and graduation, is the data source for this analysis, we can only examine college attendance and graduation, not application and admission.¹⁰ All of our outcomes are based on time windows since high school graduation. For example, five-year college graduation reflects

¹⁰We discuss application and admission more in Section 6.

college completion within five years of high school graduation (six years after the summer of the programs), not five years since college entrance.

5.1 College Enrollment and Persistence

The STEM summer programs induce a small, positive increase in college enrollment, and large shifts in institution type. The programs shift students to the HI and other elite institutions. As time goes on, we observe that students offered a seat in any program are more likely to remain enrolled throughout college and to graduate. These increases in enrollment and graduation are concentrated at the HI and other elite institutions, indicating that the programs induce a college upgrade, regardless of whether students attend the HI.

Figure 3 shows attendance at and graduation from four-year institutions, split between the HI (darker bars) and other four-year institutions (lighter bars), with a panel for each program. Table 2 shows the same estimates for college attendance as well as information about attendance at elite institutions, as designated by Barron’s college ratings. Table 3 shows the estimates for graduating college in four-, five-, and six-years. Because the time horizon differs for the cohorts, the graduation outcomes have different samples: we observe four-year graduation for all cohorts, five-year graduation for the 2014 and 2015 cohorts, and six-year graduation for the 2014 cohort alone. For four-year colleges and Barron’s most competitive colleges, we show those groups both with and without the HI included.

Almost all students in the study sample attend four-year college immediately after high school graduation (one academic year after the program summer), with 87 percent of the control group enrolling in the first year (Column 2 of Table 2). The remaining students either enroll in two-year institutions, join the military, or work. Attending one of the STEM summer programs has positive, but not statistically significant, impacts on attendance (Column 2). Column 3 shows attendance at four-year institutions other than the HI and demonstrates the programs draw students to the HI from other institutions. For the HI in particular (Column 1), 8 percent of the control group enroll. The offer of the six-week program increases enrollment by 17 percentage points, with 25 percent of students in this group enrolling. The one-week and online programs also increase HI enrollment, by

5 and 4 percentage points, respectively. The estimate for the one-week program is not statistically significant, despite being slightly larger than that of the online group.

The increases in immediate enrollment and shifts in institution reflect shifts to high-quality colleges. The HI is an elite institution, and all three programs shift enrollment there, with the largest gains for the six-week program (Column 4). Enrollment at any of Barron's most competitive institution, including the HI, sees a statistically significant gain of 17 percentage points for the six-week program, a 14 percentage point gain for the one-week program, and a 10 percentage point gain for the online program, as shown in Columns 4 and 5. For the six-week program, the institutional upgrade offered by the STEM summer program operates through the HI, as there is no difference in initial enrollment at other highly ranked institutions besides the HI (Column 5). However, the one-week and online programs increase enrollment both at the HI (by 4 to 5 percentage points, Column 1) *and* other non-HI institutions rated as most competitive by Barron's (by 6 to 8 percentage points, Column 5). Thus, for the one-week and online programs, institutional upgrades are split between the HI and other elite institutions. However, these enrollment differences are generally not statistically significant. Effects on attendance in the second and third years of college are generally similar to the initial enrollment effects (Panels B and C of Table 2). However, by the fourth year of college, the STEM summer programs show an even greater edge.

In the fourth year after high school graduation, there are positive, statistically significant impacts on enrollment in any four-year college (Column 2). Students assigned to a summer program maintain enrollment in their fourth year, whereas control group students are less likely to be enrolled. The differences in enrollment reflect a combination of control group students dropping out, taking time off, or delaying. In the first year of college, 87 percent of control group students were enrolled; by the fourth year control group enrollment falls to 75 percent. The drop for control students is primarily coming from non-HI institutions: in the third year, 74 percent of control group students are enrolled in a non-HI, four-year institution, but by the fourth year this drops to 68 percent. Persistence in the fourth year is particularly high at the HI for all summer programs (Column 2). For non-HI institutions, the one-week program induces a 6 percentage point increase in attendance in the fourth year (Column 3) and a 12 percentage point increase at Barron's most

competitive colleges besides the HI (Column 5). Additionally, the six-week program also appears to support persistence into the fourth year at non-HI institutions, as the negative impacts on attendance get smaller in the fourth year relative to earlier years (Column 3 and Column 5) but there is not a commensurate increase at the HI (which could indicate transfers across schools as opposed to differences relative to the control group in overall enrollment). By sustaining persistence through the first four years of college, the STEM summer programs set the stage for college graduation.

5.2 College Graduation

Only 53.2 percent of control group students graduate from a four-year college in four years (Table 3, Column 2). This is higher than the national four-year graduation rate observed at all U.S. institutions of 45.3 percent (U.S. Department of Education, 2020), but is low considering the students in this sample have near-perfect GPAs and standardized test scores two standard deviations above the country-wide mean. Additionally, the schoolwide four-year graduation rate at the HI for the same cohort is 87 percent. Both the six-week and one-week programs increase the likelihood that a student graduates from any institution within four years by about 8 percentage points. For the six-week program, much of the increase in four-year graduation comes from the HI, but the gains for the one-week program are split fairly evenly between the HI and other institutions. The online program increases graduation from the HI by 3.3 percentage points, with a drop at other institutions of 1.6 percentage points for an overall small gain. Column 4 of Table 3 shows that there is an overall increase in graduating with a bachelor's degree in four years from any highly-ranked institution including the HI by 5 to 12 percentage points, though this is imprecisely measured.

For five-year graduation, the STEM summer programs maintain or increase their gains in comparison to the control group (however, these estimates reflect one less cohort). Control group graduation rates increase over time, but all three programs make it more likely that a student graduates with a degree within five years; and, in the case of the one-week and online programs, these gains are driven by non-HI institutions, especially highly-ranked institutions (Column 4). We show six-year graduation impacts but note these findings rely on a single cohort. The one-week and online programs continue to have an edge in six-year graduation, again driven by non-HI,

highly-ranked institutions, but the six-week program shifts where (and when) students get their degree, not whether they do so. However, the six-year graduation rates are only available for the 2014 cohort, so some of these findings may be due to idiosyncratic differences between cohorts.¹¹

Online Appendix Table B.5 further explores where gains in graduation come from, showing four-year graduation from various college types. Most of the graduation gains come from students who would not have otherwise graduated from a four-year college on time (Column 1), while some reflect a shift from graduation at slightly lower ranked institutions (Barron’s rankings of highly competitive and very competitive, as shown in Columns 4 and 5). Thus college graduation gains are due both to a greater number of completions, as well as from upgrading institutional quality, either by shifting students to the highly ranked HI or other high-ranking institutions.

As a whole, assignment to any of the three programs increases enrollment in and graduation from a four-year college. The gains are concentrated at elite institutions: the HI for the six-week program, and the HI and other highly-ranked colleges for the one-week and online programs. The largest increases in on-time graduation come from the six-week program. The programs were successful 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 induce an increase in STEM degrees attainment.

5.3 STEM degree attainment and potential earnings

The STEM summer programs increased college graduation, both at the HI and other elite institutions. One of the stated goals of the programs is to increase the proportion of URM students in STEM careers, and a key part of the STEM pipeline is completing a major in a STEM field. We examine STEM degree completion in Table 4. This table shows overall degree attainment and then divides that by major between STEM degrees, non-STEM degrees, and bachelor’s degrees with no reported information on majors in the NSC. The table shows completion within four years in

¹¹Online Appendix Table B.2 shows graduation from the HI, limited by cohorts, and gives some credence to the idea that the differences in six-year graduation are due to fluctuations across cohorts. Current graduation outcomes include students who graduated in 2020 and 2021, the latter of whom may have been affected by the COVID-19 pandemic. We will continue to update graduation information as time goes on. See Figure 1 for a timeline of the cohorts and their on-time progress through college.

Columns 1 through 4, and within five years (one less cohort) in Columns 5 through 8. The STEM degree category includes any degree within the broad categorizations used by IPEDS that denote STEM fields, following the National Center for Education Statistics: Computer and Information Science, Engineering, Engineering Technologies, Biological and Biomedical Sciences, Mathematics and Statistics, Physical Sciences, and Science Technologies.¹² We categorize all other degrees that report a major as non-STEM degrees and also separately report degrees with no major code as “missing major.”

In 2019, 36 percent of bachelor’s degrees in the US were in STEM fields (National Science Board, 2022). The study population is predisposed to STEM fields to a greater extent than other students in the US: 69 percent of degrees earned within four years in the *control group* were in STEM fields. Even so, the STEM summer programs increased the prevalence of STEM degrees.

The six-week program increases overall receipt of four-year degrees by 8.2 percentage points: this reflects both small shifts away from non-STEM degrees (-2.6 percentage points) and unreported majors (-1.9 percentage points), and a large increase in STEM degree attainment (12.7 percent).¹³ The shift to STEM induced by the six-week program is entirely driven through the HI. For the one-week program, STEM gains come mostly from increased graduation, with a small amount of switches in major, though the increases are not statistically significant. About 39 percent ($\frac{0.035}{0.090}$) of the 9 percentage point increase in STEM degree attainment from the one-week program comes from non-HI institutions. For the online program, all of the attainment gains are concentrated in STEM fields at the HI.

The shift to STEM is even larger when looking at five-year STEM degree attainment in Columns 5 through 8. With five-year graduation, the six-week program increases STEM degrees by 20 percentage points, the one-week program by 15 percentage points, and the online program by 5 percentage points. HI-driven increases are about the same for four-year and five-year degree outcomes, meaning that the five-year gains are coming from non-HI institutions (possibly because

¹²If a student is missing information on major, but their degree is designated a bachelor of science degree (or some variation thereof, like B.S.) rather than a bachelor of arts or B.A. degree, then we count it as a STEM degree.

¹³Online Appendix Tables B.6 and B.7 report the same results with missing majors counted as non-STEM and STEM fields, respectively. The coding of missing degrees does not affect the overall finding that the programs increase STEM degree attainment, though the various coding schemes do change the magnitudes and statistical significance of some coefficients. However, the pattern of changes is not systematically smaller or larger.

a large share of students at the HI graduate on time). Online Appendix Table B.8 separates the STEM majors into specific fields, showing that engineering dominates the increase in STEM. These programs induce not just college graduation, but college graduation in STEM fields.

The shift in STEM degrees is due to degree completion, not differences in STEM interest for applicants offered a seat in the program. Turning briefly to survey evidence, the programs do not change reported interest in STEM majors immediately after the program (Figure 6), with 93 percent of both treatment and control students finishing the program summer reporting plans to major in a STEM degree. Once in college, both treatment and control students maintain very high levels of interest in STEM degrees, with about 83 percent of students planning to declare a STEM major. There are some differences when it comes to intentions to pursue a STEM career: both the six-week and online programs increase the likelihood that students report wanting to pursue a STEM career in the fall of the senior year of high school, and the six- and one-week programs induce similar gains in STEM career intentions mid-college (though they are not statistically significant). Thus, we conclude that the increase in STEM degree attainment is due to the groups with program offers being more likely to follow through on STEM intentions—perhaps due to an upgrade in institutional quality that we discuss more below—rather than the programs inspiring a greater degree of interest in STEM fields.

To understand how the shift to STEM may influence future earnings, we use the concept of “potential earnings” from Sloane et al. (2021). Sloane et al. (2021) define potential earnings as the earnings of “the median middle-aged, US-born, White male” by major, and calculate this using the American Community Survey from 2014 to 2017. This measure represents the major-specific potential earnings subtracting out labor market experiences such as discrimination and penalties for taking time off for family responsibilities. We assign to each student in the sample the potential earnings of their major (in natural log units) and plot the distribution of majors by earnings, for those who graduate with a degree in four years and in five years (Figure 4). The changes in distribution of degrees will thus reflect changes in potential earnings due to the composition of majors, *not* differences due to the increased graduation rate induced by the program. Thus it is perhaps a conservative estimate of impacts on potential earnings. Included in the figure are the

program-specific impacts on potential earnings based on graduates. Online Appendix Table B.9 reports these regression estimates, along with estimates that impute earnings for non-graduates.

Figure 4 shows the same shift to engineering shown in Online Appendix Table B.8 but highlights that engineering is the major with the highest potential earnings. The figure also shows a decline in psychology majors for the six- and one-week programs. The average difference in potential earnings is about 4 percent for the six- and one-week programs, and 1.4 percent for the online program (not significant) when using four-year graduation. With five-year graduation, these estimates are even larger, at 6 percent for the six-week program, 5 percent for the one-week program, and 2 percent for the online program. Although these estimates cannot predict exact earnings, they show that those offered seats in the program are set up to enter fields where their earnings are at least 2 to 6 percent higher than they would have been in the absence of the STEM summer programs.

5.4 Heterogeneity

In Online Appendix Tables B.10 through B.13, 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 as these groups include populations of particular interest for increasing representation in STEM fields. Both male and female students appear to benefit equally from the six-week program, but female students have slightly larger gains from the one-week and online programs, especially for graduating from a highly ranked institution (Online Appendix Table B.10). Larger gains for female students are perhaps not surprising 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.

In Online Appendix Table B.11, we show estimates for Black, Hispanic, and Native American individuals separately from students who do not identify 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, 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 almost all of the benefits of the online program accrue for URM students. Again, the sample of non-URM students is quite small, and there are large and meaningful benefits for both groups.

The clearest difference comes from a comparison of students who report receiving subsidized lunch at their high school with those that do not (Online Appendix Table B.12). Although exposure to the programs is generally beneficial for both groups, students who do *not* receive subsidized lunch reap larger gains, especially for graduating from an elite institution and graduating with a STEM degree. Students with more resources may be better poised for an institutional upgrade, perhaps because of fewer concerns about college costs. Increasing representation by race in STEM fields will not necessarily increase representation by family income, and vice versa.

Examining differences by first-generation college-going status (defined as no parent ever attending a four-year college) shows that both groups benefit from participation in a STEM summer program, but that the six-week program is more effective for first-generation students than non-first generation students, likely due to the large shift to the HI. Gains at elite institutions from the one-week and online programs (which we showed earlier were split between the HI and other elite institutions) are more consistent for the non-first-generation group. The programs may be most effective for first-generation students who are directly influenced to attend the HI, and the non-HI benefits of the programs may accrue more for students who already have social capital.

5.5 Validating the Experiment

This section considers several tests of whether the estimation strategy under constrained random assignment effectively accounts for selection bias and the extent to which heterogeneous treatment effects account for our findings, the role of cohort variability, and the importance of limiting the range of randomization scenarios. We compare estimates from Equation 1 to alternative specifications on key outcome variables at the HI and across institutions.

5.5.1 Investigating selection bias

We argue that our conditional random assignment specification is as good as random because the conditions of random assignment are based on observable characteristics that we fully account

for with randomization blocks. To make this explicit, in Online Appendix Tables A.3 and A.4, we show how controlling for randomization block effectively controls for selection bias. Panel A of these tables first shows the main specification, but *without* controls for randomization block (retaining controls for cohort, gender, and geographic preference). In this case, most of the estimates are biased upward in comparison to the main estimates with the blocks (Panel B) because students with higher rating variables have greater success in college, regardless of program participation. Panel C adds a control for the rating variable to the main estimates in Panel B, which leaves the findings unaffected. This is not surprising because within the block, the rating variable is randomly assigned. Panel D shows an alternative way of controlling for assignment: removing the rank-component of the blocks as in Panel A, but controlling for rating variable as in Panel C. In this case, the estimates are of slightly smaller magnitudes at the HI but tell the same story as the main estimates. Outside of the HI, the reduction in magnitudes is slightly larger but again show positive effects of program offer. Any of the estimates in Panels B through D are causal estimates of the programs' effectiveness, and they all account for the selection bias shown in Panel A. We use the estimates in Panel B as our main specification, as this was the intended design of our experimental estimates. In Online Appendix A.1.6, we also discuss the scope for heterogeneous treatment effects in this modified random assignment to explain the treatment effects. Although the heterogeneous response of highly rated students may contribute to the magnitudes for the one-week program, the evidence that differential response by rating drives program effects is limited (Online Appendix Tables A.11 and A.12). We also show in this appendix that estimates of the online program effect, both within block and across blocks—using the alternative controls from Online Appendix Table A.3—show very similar treatment effects for the online program, demonstrating that our “link” across blocks is sound.

5.5.2 Alternative specifications

Online Appendix Tables A.5 and A.6 report alternative specifications for outcomes across institutions and at the HI, respectively. Panel A shows the main estimates without baseline covariates. As expected, precision decreases a bit, but the magnitude and direction of the estimates remain

the same. Panels B through D remove each cohort in turn. We expect there to be idiosyncrasies across cohorts due to sampling variation, but each panel reports estimates that are generally in line with the main findings. In some cases there is a loss of precision due to smaller samples. Panel B is notable in that it removes the 2014 cohort. This cohort had the most modifications to random assignment, with more blocks than in the two subsequent cohorts. If the study design with a greater number of blocks did not remove selection bias as fully as the design with fewer blocks, we would expect that the estimates in Panel B would be *smaller* than the main findings because removing 2014 would remove upwardly biased estimates. Instead, the results limited to the two later cohorts with a design closer to complete random assignment show either very similar or larger impact estimates.

5.5.3 Randomization inference

A different, but related, threat to validity comes from the randomization scheme. By imposing randomization strata and modifying complete random assignment, the study limits the range of potential outcomes that is possible. For example, if a state is preferred in the assignment for representational reasons, it is always going to 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 using randomization inference. Specifically, we re-randomize applicants to the programs, using the same randomization criteria (blocks, location, gender preferences, etc.) 1,000 times, and compare the estimate from our main specification to the distribution of estimates from the 1,000 randomizations. Each randomization faces the constraints imposed by our research design, so we are limiting possible comparisons to those that might actually occur rather than a hypothetical scenario of full randomization.

We display results from this exercise in Online Appendix Figures B.3 and B.5. Each panel shows the distribution of treatment estimates from the 1,000 hypothetical randomizations (bars), compared to the main specification estimate (dashed line). If the impact estimate from the main specification is at or above the 97.5th percentile (two-tailed test) or 95th percentile (one-tailed test), then that implies statistical significance. For any college attendance, impact estimates are at

the 76th to the 83rd percentiles, but at elite institutions, impacts on attendance are at the 97th to 99th percentile—generally similar to the pattern of findings in the main estimates. Impacts on graduation and STEM degree attainment range from the 80th percentile to the 98th percentile, with six-week program effects more likely to be outliers. Again, this pattern aligns with our main estimates.¹⁴ We consider the randomization inference exercise as confirmatory evidence that we can draw inferences from our constrained randomization scheme using traditional statistical methods.

6 Mechanisms

In this section, we explore mechanisms behind the successful STEM summer programs, focusing on graduation from a four-year college and obtaining a STEM degree as the main outcomes of interest. For many of these analyses, we use data from student surveys, which have some differential response by treatment arm. As we note in Section 3, we use unweighted survey responses due to the inconsistent influence of reweighting schemes to address survey response bias, and instead consider findings from the surveys suggestive rather than conclusive.

There are several reasons why the programs might improve college graduation and STEM degree attainment, and we focus on three, human capital-related explanations here.¹⁵ First, the programs may drive impacts by increasing participants’ human capital with respect to the college application process, which shifts where students go to college. In turn, this drives differential institutional quality, which causes the differences in outcomes that we observe. We also discuss the extent to which institutional upgrading is due to greater likelihood of admission because of a

¹⁴In the case of application and admission to the HI, the main specification estimate is above the 99th percentile of the distributions of placebo randomization estimates, implying that it is extremely unlikely that the estimate was due to chance. When we turn to the impacts on attendance at and graduation from the HI, the patterns again follow the main estimates closely. The estimates for the six-week program are larger than all of the estimates from the placebo randomizations. The one-week and online programs have the impact estimates at the 80th to the 87th percentiles of the placebo randomizations. However, this, too, is in line with our main estimates, as attendance and graduation impacts for the one-week program tend to be of modest magnitude and not statistically significant.

¹⁵We focus on these explanations both because they are meaningful, plausible mechanisms and because we have the most data to bring to bear on these mechanisms. Another interesting possibility is that participation in the program serves as a signal to college admissions officers. Signaling is probably most powerful at the HI, and many of the one-week and online program gains are at non-HI institutions. Additionally, while this channel may play a role, we note that it would primarily influence *admission* but not necessarily what happens during college. So, it is unlikely to be an explanation for the increase in STEM degree attainment, especially given that students in both the treatment and control groups report an intention to major in STEM at very high rates. Nonetheless, we discuss signaling with respect to college admissions when we discuss college application behavior below.

signaling effect of the programs. Second, the program may improve participants' human capital with respect to subject matter knowledge, helping them get ahead in college and more likely to graduate. Finally, we consider whether there is an increase in participants human capital with respect to their soft skills, which better position them to succeed in college. As we detail below, we have the most evidence that the first channel explains our findings, though we cannot fully rule out other explanations.

6.1 College application behavior and college quality shifts

The STEM summer programs have potential to affect the trajectories of young people, but ultimately, even the most intense six-week program is a short period in a young person's life. However, by shifting students to different colleges from those they would have attended otherwise, the programs may set students on a different path. We have shown that program participation increases enrollment in and graduation from the HI and other highly ranked colleges. Here, we show the primary path for shifting students into these institutions is through changing college application strategy. We first show that application and admissions patterns change for program participants, and then discuss how this shift induces institutional upgrading.

We observe increased application to the HI (via administrative data from the HI, displayed in Online Appendix Figure B.1 and Online Appendix Table B.1) and other elite institutions (via survey data, displayed in Table 5). About 31 percent of control group students apply to the HI. This would be a very high rate of application for typical high school students, but program applicants are a selected group who were interested in the outreach office's summer programs. Even so, being assigned to one of the summer programs more than doubles application to the HI for all the programs, with almost all students (78 percent) in the six-week group applying to the HI.

Almost 11 percent of the control group are admitted to the HI, again demonstrating that the study sample is a selected group: the HI typically admits fewer than 10 percent of applicants. Given that a little over 30 percent of the control group applied, this is an admission rate of about 34 percent. Admissions rates are generally on par with the control group admission rate among those with program offers. About 31 percent of students assigned to the six-week program are

admitted to the HI, with an implied admission rate of 40 percent. For the one-week program, about 21 percent are admitted, with an admission rate of about 30 percent; for the online program, 19 percent are admitted, with an admission rate of about 29 percent. Thus, while there is a small bump in likelihood of admission for the six-week program, the main difference between the control group and treatment groups is likelihood of applying. Students offered seats at the STEM summer programs are more likely to be admitted to the HI, but for the most part this seems to be due to greater likelihood of application, rather than greater admission conditional on application. We explore this in more detail below.

Enrollment and graduation changed beyond the HI, with programs shifting students to elite institutions other than the HI (especially the one-week and online programs), so admission behavior to these institutions likely changed as well. We turn to survey responses to examine changes in application behavior more broadly. Survey respondents reported their college applications and admission in a survey conducted in May of their senior year of high school. As we show in Table 5, those offered seats at STEM summer program are more likely to apply and be admitted to a Barron's most competitive institution, even when excluding the HI from that category (Column 5). The one-week and online programs also induce 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 comes from a reduction in application to a single school (Column 4). Attending the six-week program eliminates the (small) possibility that students apply to only one school, and attendance at the one-week or online program halves this likelihood. As shown by enrollment in higher-quality institutions (Table 2), these changes in application behavior translate into more chances to attend a high-quality institution.

These improvements in college application come from increased knowledge about the college admissions process. The programs also improve students' college resources through knowledge of both the landscape of available colleges and the college application process itself, as we show in Online Appendix Table C.3. Survey responses from the fall of the senior year of high school show that the programs—especially the six-week program—increase sources of college application advice (Columns 1 through 4), with students assigned to the treatment group reporting greater likelihood

of getting advice from friends and a teacher or counselor. Assignment to a program also increased familiarity with non-HI institutions: students were more likely to report familiarity with a technical institution (by 7 to 10 percentage points), an elite institution (by 1 percentage point), and a liberal arts college (by 7 to 11 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 translate into increased enrollment in high-quality institutions, as we show in Section 5. We investigate whether enrollment in these institutions spurs the college graduation and STEM results we observe in Figure 5, which compares the *institution-level* outcome of the university attended immediately after high school graduation to the individual-level outcome. To generate the institution outcomes, we assign the institution-specific graduation rate to those who attend an institution in their first year of college.¹⁶ This outcome then measures the “predicted” graduation rates for each treatment group, if we assume that students in the study group graduate at the graduation rate of the institution they attend. We construct a similar outcome for STEM degrees, which is the institution specific share of degrees that are in STEM fields, based on degrees reported to IPEDS.

The results from this exercise are in Figure 5, with more details in Online Appendix Table B.16. Panel A compares the difference in “actual” graduation with the “predicted” graduation rates. For the six-week and one-week programs, the change in actual four-year graduation rates almost exactly parallels the difference in predicted graduation rates (for the online program, the change in graduation is about one-third the size of the difference in expected graduation), which is consistent with many of the program benefits operating through students’ upgrading college quality. When we conduct a similar analysis for STEM degree share, we see in Panel B that the six-week and online programs increase attendance at institutions with higher proportions of STEM degrees, but the actual gains in STEM degree receipt are even greater than the predicted difference. Thus, students offered STEM summer programs increase their likelihood of graduation by the amount predicted due to institutional upgrading, but they increase their STEM degree attainment to an

¹⁶For students who attend community colleges or do not attend college, we substitute zeroes instead.

even greater extent, providing evidence that program participation helps students achieve their intention to obtain a STEM degree in very competitive institutions.

6.2 Signaling

We examine how much of the application effect is due to applying as opposed to greater odds of admission upon applying to colleges in Online Appendix Table B.14. In this table, we show application, admission, and admission conditional on applying to the HI (Columns 1 through 3), at any Barron's most competitive institutions (Columns 4 through 6), and at Barron's most competitive institutions excluding the HI. The latter two categories are restricted to the survey respondents who report admissions and enrollment. The conditional admissions estimates show the increased likelihood of being admitted to an institution for the STEM summer programs, which may be due to signaling or program-induced improvements in students' application packages or human capital. At the HI, part of the admissions gains are due simply to increased likelihood of application, but part of the gains are from an additional bump in admissions for the STEM summer programs (15 percentage points for the six-week program, and 5 to 5.5 percentage points for the one-week and online programs, though these differences are not statistically significant). At non-HI elite institutions, the online and one-week programs increase the likelihood of conditional admissions by 5 and 6 percentage points, respectively, though only the former is statistically significant. Overall, these findings show that there are some increases in conditional admission but the role of changes in application behavior is large.

Panel B compares the HI STEM summer programs to a restricted control group, where control students are limited to those who report attending any STEM summer experience.¹⁷ This control group is more likely to be female and has slightly higher rating scores than control students who do not participate in STEM summer experiences, but is otherwise quite similar to the control group overall (see Online Appendix Table B.18). If we assume different STEM summer experiences impart similar human capital gains, comparing those assigned to the HI STEM summer programs with those in the control group with STEM summer experiences tests whether the HI programs have

¹⁷These include similar programs at other colleges and universities, nonprofit programs like Girls Who Code, internships in STEM or medical fields, and taking STEM courses.

signaling value above and beyond that of having any STEM summer programming. Even with this comparison group, there is still a conditional admission gain of 10 percentage points at the HI for the six-week program, but there is no difference for the one-week and online programs. We interpret this to mean there is likely some role for signaling for the six-week program at the HI. However, outside the HI, there appears to be little difference in conditional admissions between the HI STEM summer groups and control group members with STEM summer experiences, implying that outside of the HI there are no special gains for HI STEM summer program compared to any STEM summer programs who operate outside the application channel.¹⁸ These estimates are consistent with some role for signaling, but do not prove its existence. Even when comparing to control students with STEM summer experiences, our assumption about similar human capital gains across programs could be wrong, and the additional boost for the six-week program compared to control students with STEM summer experiences might be due to human capital gains from the program. These findings provide a ceiling on the signaling mechanism and confirm the application channel accounts for many of the program effects.

6.3 Subject matter knowledge, skills, and confidence

Another reason the STEM summer programs may lead to college graduation and STEM degrees is by increasing students' STEM subject matter knowledge, helping them get ahead in college. This could occur in two ways: the programs could change high school classes taken prior to college entrance, or they could directly increase knowledge of STEM, better preparing them for future STEM majors. Additionally, the programs may improve more general skills such as study skills, as well as soft skills such as confidence and self-esteem.

Table 6 shows impacts on both high school course plans and on a direct measure of human

¹⁸We also compare graduation and STEM degree outcomes between those assigned to HI STEM summer programs and the selected control group with STEM summer experiences in Online Appendix Table B.15). This comparison contrasts randomly assigned students with a group that will both have gains from their STEM summer experiences and a selection effect, as these students sought out additional STEM experiences. Thus, we expect program impacts to be smaller when comparing to this group, and they are, as seen in Panel B. However, we note that the HI STEM programs still maintain an increase above this selected comparison group in five-year graduation (Column 6), due to increased five-year graduation at the HI for the six-week program (Column 2) and increased five-year graduation from other institutions for the one-week and online programs (Column 10), though the magnitudes are smaller and they are not statistically significant with the full control group (Panel A). In terms of STEM degrees, the shift is less stark, which makes sense if we think the control group represents the most STEM-ambitious comparison.

capital. Most students were already planning to take at least one Advanced Placement (AP) or International Baccalaureate (IB) course in their senior year in high school (control group mean of 89 percent), and program participation had positive impacts on these intentions (only statistically significant for the one-week program), as shown in Panel A. When it came to increasing AP/IB coursework, the programs increased computer science take-up. About 10 percent of the control group planned to enroll in such a course, and program offers increased that by 7 to 10 percentage points. Perhaps because science and math advanced coursework was already extremely popular (control means of 73 to 74 percent for both), the programs made little difference in science and math advanced coursework. During the fall of the senior year of high school, applicants offered seats in the programs were slightly better able to answer a calculus question, though the differences were not statistically significant (Column 7 in Panel B). The minor calculus improvement and coursework changes in non-core subjects indicate that although the programs may impart STEM knowledge and inspire subsequent high school coursework, these changes are relatively small.

We illustrate program impacts on soft skills in Table 6, which shows student survey responses to questions about life skills, study skills, confidence, interest in learning, and attention span.¹⁹ 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 be their first time away from their family, and we observe an increase in the self-reported life skills that one would gain in such a situation. As predicted, gains in life skills are larger for residential programs as opposed to the online program (even though this program had a short campus visit). We also see that the programs increase self-reported study skills, such as asking questions and taking notes. However, there is no statistically significant change in confidence, which included a self-assessment of students' math ability, though the direction is positive. The negative coefficients for attention span (statistically significant for the one-week and online programs) may reflect students engaging with more challenging material over the summer and in the fall of their senior year of high school. The positive, statistically significant increases in life and study skills, and the positive but not statistically significant increases in confidence and enjoyment of intellectual activities are all consistent with the idea that the STEM

¹⁹See Online Appendix C for details on the variables included and how we generated indices from individual variables. Online Appendix C also includes additional survey responses not discussed in the text.

summer programs better prepare students to succeed in and graduate college and may contribute to the increases in STEM degree attainment that we cannot account for solely through upgrades in institutional quality.

7 Conclusion

We present evidence from the first randomized controlled trial to investigate the impact of STEM summer programs for underrepresented youth, examining a suite of programs that includes a six-week program, a one-week program, and an online program. Program offers increase student matriculation at the HI and other elite universities. Students exposed to the STEM summer programs are more likely to persist through college, with the most notable difference coming in the fourth year of college, when a larger share of control group students are no longer enrolled. The programs increase overall four-year college graduation, with gains concentrated at the HI and other elite institutions. The programs also induce increases in attainment of STEM degrees. Response to program assignment, especially at the HI, is largest for the most intensive, six-week program, but both the one-week and online programs consistently increase graduation and STEM degree attainment. The shift in major toward STEM degrees may increase potential earnings by 2 to 6 percent. We show evidence consistent with the idea that a change in college application behavior shifts students to higher-quality institutions, which drives the gains in college graduation.

The benefits for generating STEM degrees, in contrast, go beyond what we would expect solely based on institutional quality (as measured by share of STEM degrees on campus). This is a highly STEM-motivated group, with almost all of the study sample intending to major in STEM in college. Like students across the nation, those intentions face a “leaky pipeline” where students leave STEM preparation throughout college. The STEM summer programs make students more resistant to leaks in the pipeline, though there are still STEM-intending students who do not ultimately complete a STEM degree. We cannot definitively say why this is but note that the programs have 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 and 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 program generates the largest response, but is also the most expensive, in contrast to the one-week and online programs. In 2015, the six-week program cost about \$15,000 per student, while the one-week and online programs cost about \$2,000 per student. Perspectives on which program is the most cost-effective will differ by the objective. For the HI, the six-week program produces an increase in on-time graduation four times as large as the one-week or online program, using the treatment-on-the-treated estimates from Online Appendix Table B.4. But if a policymaker is interested in overall graduation from the most competitive institutions, the six-week program only outperforms the one-week program by a small amount (an increase of 13 percentage points versus an increase of 12 percentage points, as shown in Online Appendix Table B.3), though it still doubles the effect of the online program. The comparison is similar for on-time STEM degrees. A more detailed cost-effectiveness analysis is left as a future exercise.

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 benefit from a more diverse and inclusive STEM classroom. As the U.S. Supreme Court continues to erode affirmative action as a component of higher education admissions, more colleges and universities may turn to programs like 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 to higher education, not earlier in the pipeline.²⁰ Our findings show that focusing on higher education after students apply to college may miss a key opportunity to intervene in students lives *before* they apply to college—the point in time crucial to the institutional choices that may ultimately help students succeed.

²⁰In 2011, there were over 250 federal programs and \$3.4 billion invested in the STEM pipeline (Granovskiy, 2018), including over \$1 billion in funding through the National Science Foundation (NSF) with the specific goals of increasing diversity and representation (see NSF budgets for “Broadening Participation” efforts here: https://www.nsf.gov/od/broadeningparticipation/bp_investments.jsp.) About three-quarters of federal investment supports undergraduate and graduate education and training, with less support for K-12 initiatives. (Authors’ calculations from Granovskiy (2018).)

References

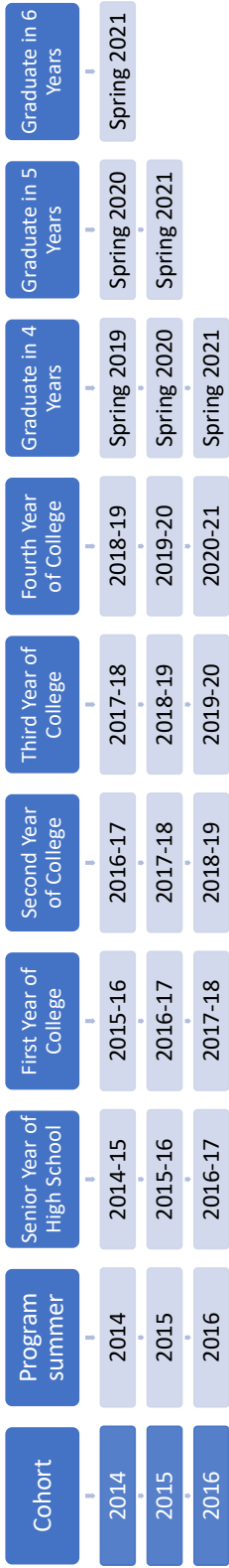
- Altonji, J. G., P. Arcidiacono, and A. Maurel (2016, January). The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 5. Elsevier.
- Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics* 4, 185–223.
- Anderson, M. 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, R. J., S. A. Imberman, and M. F. Lovenheim (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review* 74, 101923.
- Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics* 121(1-2), 343–375.
- Arcidiacono, P., E. M. Aucejo, and V. J. Hotz (2016). University differences in the graduation of minorities in stem fields: Evidence from california. *American Economics Review* 106(3), 525–562.
- Arcidiacono, P. and M. Lovenheim (2016). Affirmative Action and the Quality-Fit Trade-Off. *Journal of Economic Literature* 54(1), 3–51.
- Astorne-Figari, C. and J. 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, S. and G. W. Imbens (2017). The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, Volume 1, pp. 73–140. Elsevier.
- Avery, C. (2010, September). The effect of college counseling on high-achieving, low-income students. Working Paper 16359, National Bureau of Economic Research.
- Avery, C. (2013, October). Evaluation of the college possible program: Results from a randomized controlled trial. Working Paper 19562, National Bureau of Economic Research.
- Bailey, M. J. and S. M. Dynarski (2011). Gains and gaps: Changing inequality in us college entry and completion. Technical report, National Bureau of Economic Research.
- Baker, R., D. Klasik, and S. F. Reardon (2018). Race and stratification in college enrollment over time. *AERA Open* 4(1), 2332858417751896.
- Becker, C. M., C. E. Rouse, and M. 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, E. (2010). To be or not to be: Major choices in budding scientists. In *American universities in a global market*, pp. 69–98. University of Chicago Press.
- Bettinger, E. P. and B. T. Long (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review* 95(2), 152–157.
- Bleemer, Z. (2021). Top percent policies and the return to postsecondary selectivity.
- Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after california’s proposition 209. *The Quarterly Journal of Economics* 137(1), 115–160.
- Bradford, B. C., M. E. Beier, and F. L. Oswald (2021). A meta-analysis of university stem summer bridge program effectiveness. *CBE Life Sciences Education* 20(2), ar21.
- Carrell, S. and B. Sacerdote (2017). Why do college-going interventions work? *American Economic Journal: Applied Economics* 9(3), 124–51.
- Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly journal of economics* 125(3), 1101–1144.
- Castleman, B. and J. Goodman (2018, 01). Intensive College Counseling and the Enrollment and Persistence of Low-Income Students. *Education Finance and Policy* 13(1), 19–41.
- Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics* 135(3), 1567–1633.
- Cohodes, S. R. and J. S. Goodman (2014, October). Merit aid, college quality, and college completion: Massachusetts’ adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics* 6(4), 251–85.
- Cook, L. D., J. Gerson, and J. Kuan (2021, October). Closing the innovation gap in pink and black. Working Paper 29354, National Bureau of Economic Research.
- Cosentino, C., C. Speroni, M. Sullivan, R. Torres, et al. (2015). Impact evaluation of the rwjf summer medical and dental education program (smdep). Technical report, Mathematica Policy Research.
- Darolia, R., C. Koedel, J. B. Main, J. F. Ndashimye, and J. Yan (2020). High school course access and postsecondary stem enrollment and attainment. *Educational Evaluation and Policy Analysis* 42(1), 22–45.
- De Philippis, M. (2021). Stem graduates and secondary school curriculum: does early exposure to science matter? *Journal of Human Resources*, 1219–10624R1.
- Dillon, E. W. and J. A. Smith (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics* 35(1), 45–66.
- Dutz, D., I. Huitfeldt, S. Lacouture, M. Mogstad, A. Torgovitsky, and W. van Dijk (2021). Selection in surveys. Working Paper 29549, National Bureau of Economic Research.

- Dynarski, S., C. Libassi, K. Michelmore, and S. 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, R. W., F. Hoffmann, and P. Oreopoulos (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review* 104(8), 2567–91.
- Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men’s and women’s stem persistence. *Labour Economics* 46, 211–226.
- Gerber, T. P. and S. Y. Cheung (2008). Horizontal stratification in postsecondary education: Forms, explanations, and implications. *Annu. Rev. Sociol* 34, 299–318.
- Goodman, J., M. Hurwitz, and J. Smith (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics* 35(3), 829–867.
- Granovskiy, B. (2018). Science, technology, engineering, and mathematics (stem) education: An overview. crs report r45223, version 4. updated. *Congressional Research Service*.
- Green, A. and D. Sanderson (2018). The roots of stem achievement: An analysis of persistence and attainment in stem majors. *The American Economist* 63(1), 79–93.
- Griffith, A. 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, A. L. and J. B. Main (2019). First impressions in the classroom: How do class characteristics affect student grades and majors? *Economics of Education Review* 69, 125–137.
- Hoekstra, M. (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, F. and P. Oreopoulos (2009). A professor like me the influence of instructor gender on college achievement. *Journal of human resources* 44(2), 479–494.
- Hofstra, B., V. V. Kulkarni, S. M.-N. Galvez, B. He, D. Jurafsky, and D. A. McFarland (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences* 117(17), 9284–9291.
- Hoxby, C. and C. Avery (2013). The missing “one-offs”: The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, 1–65.
- Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The allocation of talent and us economic growth. *Econometrica* 87(5), 1439–1474.
- Joensen, J. S. and H. S. Nielsen (2016). Mathematics and gender: Heterogeneity in causes and consequences. *The Economic Journal* 126(593), 1129–1163.
- Kaganovich, M., M. Taylor, and R. Xiao (2021). Gender differences in persistence in a field of study.
- Kinsler, J. and R. Pavan (2015). The specificity of general human capital: Evidence from college major choice. *Journal of Labor Economics* 33(4), 933–972.

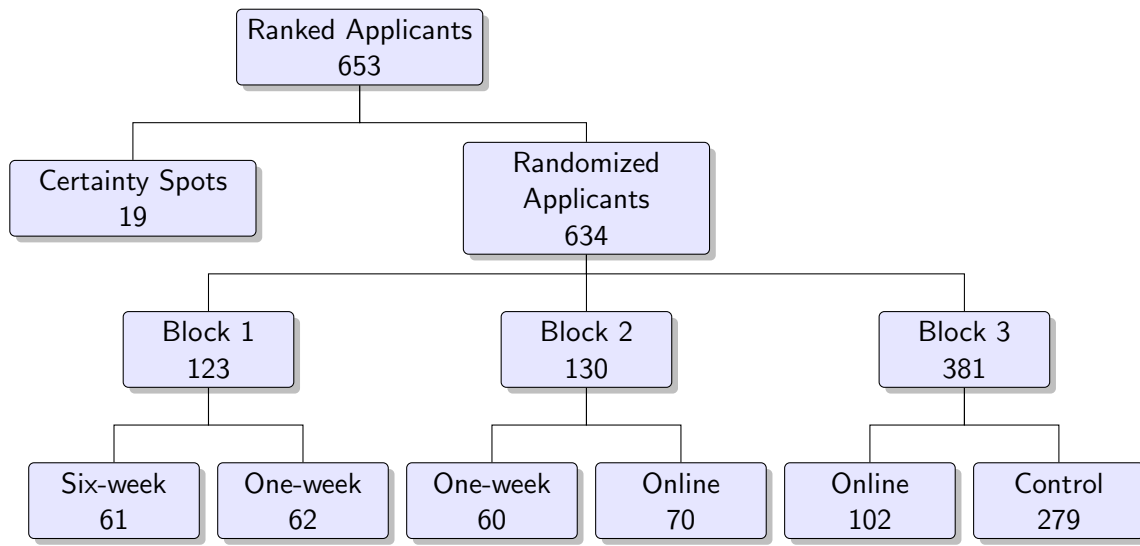
- Kitchen, J. A., P. Sadler, and G. Sonnert (2018). The impact of summer bridge programs on college students' stem career aspirations. *Journal of College Student Development* 59(6), 698–715.
- Kitchen, J. A., G. Sonnert, and P. M. Sadler (2018). The impact of college-and university-run high school summer programs on students' end of high school stem career aspirations. *Science Education* 102(3), 529–547.
- Kugler, A. D., C. H. Tinsley, and O. Ukhaneva (2021). Choice of majors: are women really different from men? *Economics of Education Review* 81, 102079.
- National Science Board (2021). The stem labor force of today: Scientists, engineers and skilled technical workers: Science and engineering indicators 2021. Technical Report NSB-2021-2, National Science Foundation, Arlington, VA.
- National Science Board (2022). Higher education in science and engineering: Science and engineering indicators 2022. Technical Report NSB-2022-3, National Science Foundation, Arlington, VA.
- Owen, S. (2020). College field specialization and beliefs about relative performance. Technical report.
- Owen, S. (2021). Ahead of the curve: Grade signals, gender, and college major choice. Technical report.
- Parrotta, P., D. Pozzoli, and M. Pytlikova (2014). The nexus between labor diversity and firm's innovation. *Journal of Population Economics* 27(2), 303–364.
- Price, G. N. (2005). The causal effects of participation in the american economic association summer minority program. *Southern Economic Journal* 72(1), 78–97.
- Price, J. (2010). The effect of instructor race and gender on student persistence in stem fields. *Economics of Education Review* 29(6), 901–910.
- Riegle-Crumb, C., B. King, and Y. Irizarry (2019). Does stem stand out? examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher* 48(3), 133–144.
- Robles, S. (2018). The impact of a stem-focused summer program on college and major choices among underserved high-achievers. (2018.03).
- Sass, T. R. (2015, January). *Understanding the STEM Pipeline. Working Paper 125*. National Center for Analysis of Longitudinal Data in Education Research.
- Sloane, C. M., E. G. Hurst, and D. A. Black (2021). College majors, occupations, and the gender wage gap. *Journal of Economic Perspectives* 35(4), 223–48.
- U.S. Department of Education (2020). Digest of education statistics 2020. table 326.10. *National Center for Education Statistics*.
- Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics* 32(4), 711–754.

Figure 1: Experiment Timeline and Available Data



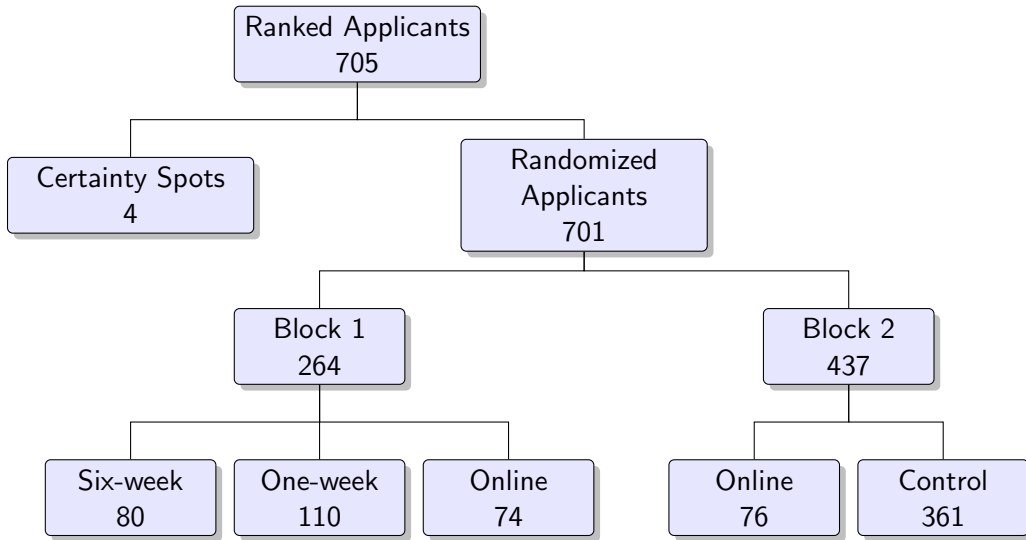
Notes: This figure shows student progress over time by experimental cohort, assuming students maintain on-time progress through college. The most recent available information on college enrollment and graduation reflects the spring 2021 semester.

Figure 2a: Randomization Design: Cohort 1 (2014)



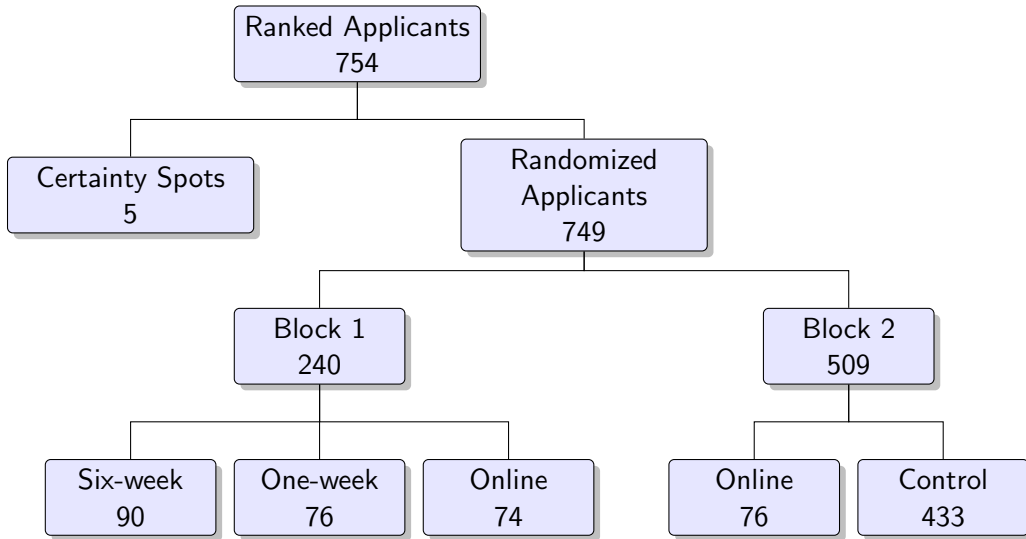
Notes: This figure shows the blocked randomization design in 2014. Certainty spots were applicants offered admission with certainty and excluded from the experimental analysis. All other ranked applicants were subject to random assignment within block. Block assignment reflects applicant ratings, math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

Figure 2b: Randomization Design: Cohort 2 (2015)



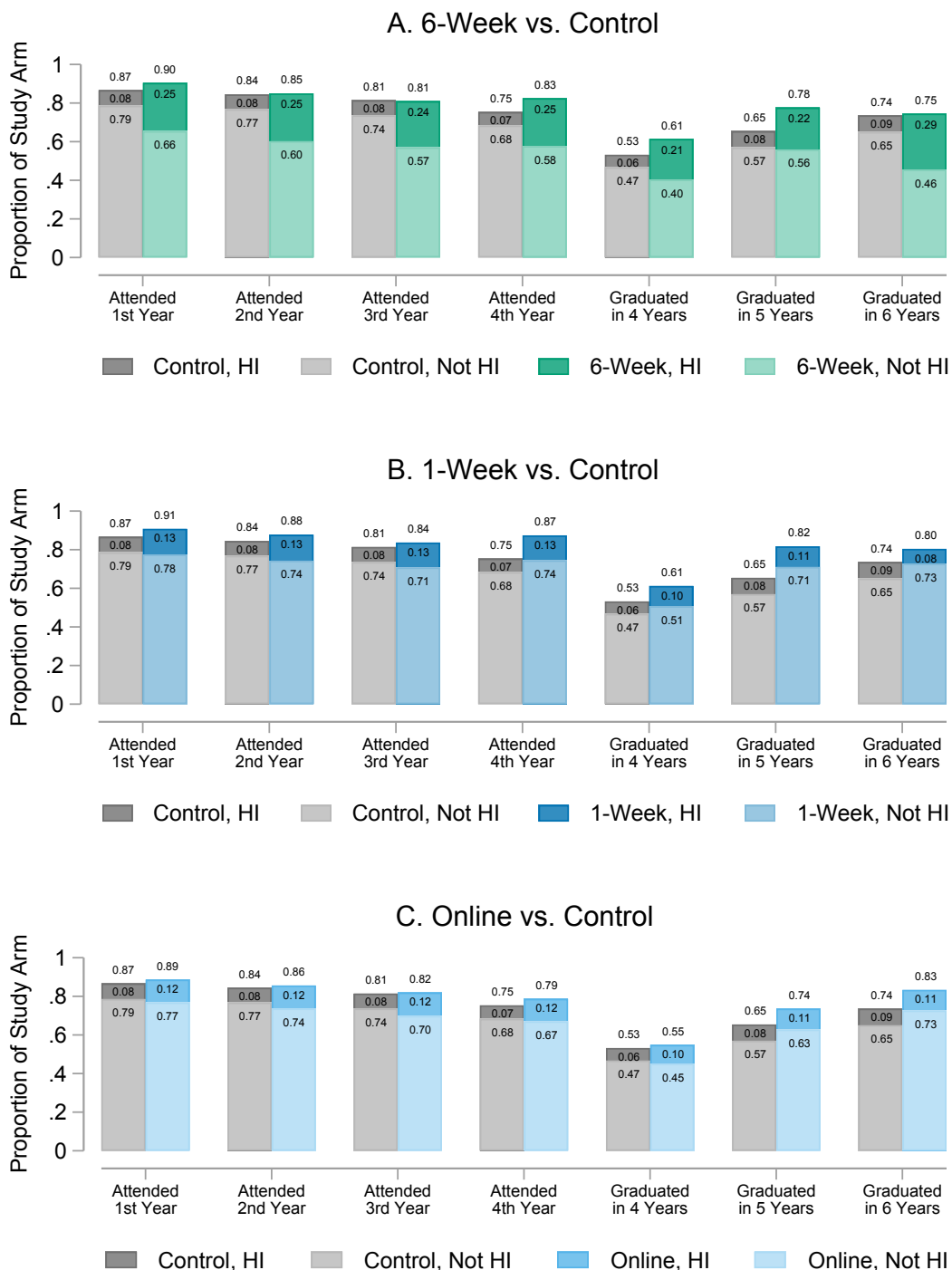
Notes: This figure shows the blocked randomization design in 2015. All other notes are the same as in Figure 2a.

Figure 2c: Randomization Design: Cohort 3 (2016)



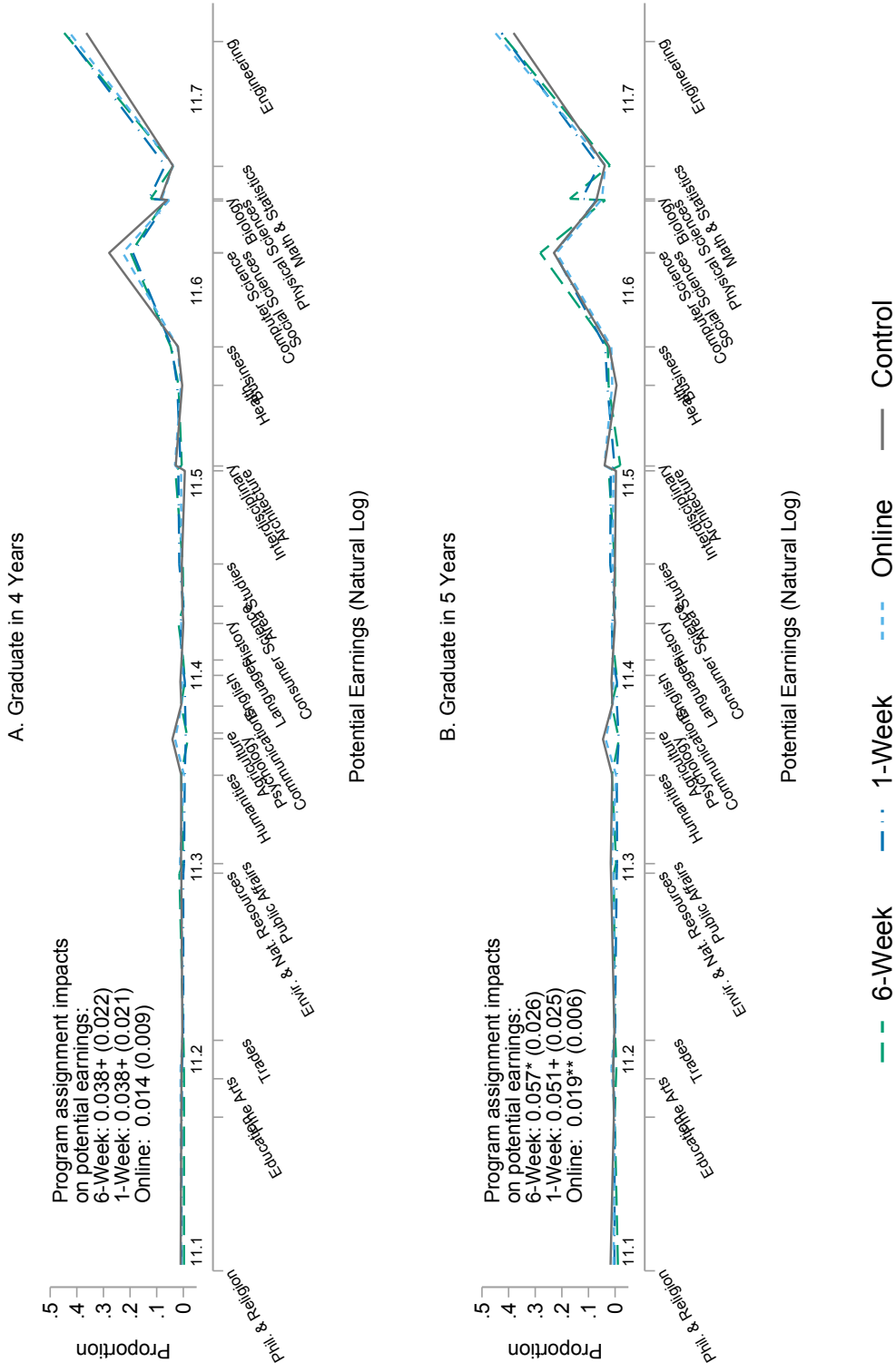
Notes: This figure shows the blocked randomization design in 2016. All other notes are the same as in Figure 2a.

Figure 3: The Impact of STEM Summer Assignment on Four-Year College Attendance and Graduation



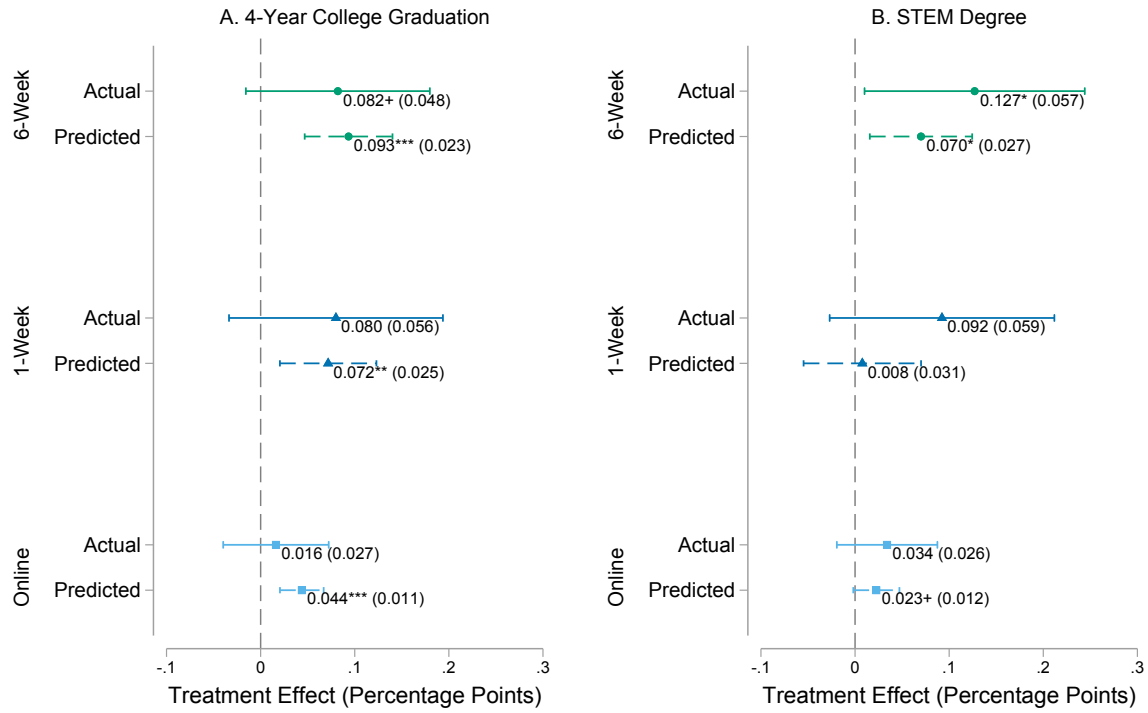
Notes: This figure summarizes impact estimates for four-year institution outcomes, with and without the HI included. For details on the specification and exact point estimates and standard errors, see Tables 2 and 3.

Figure 4: The Impact of STEM Summer Programs on Potential Earnings by Major



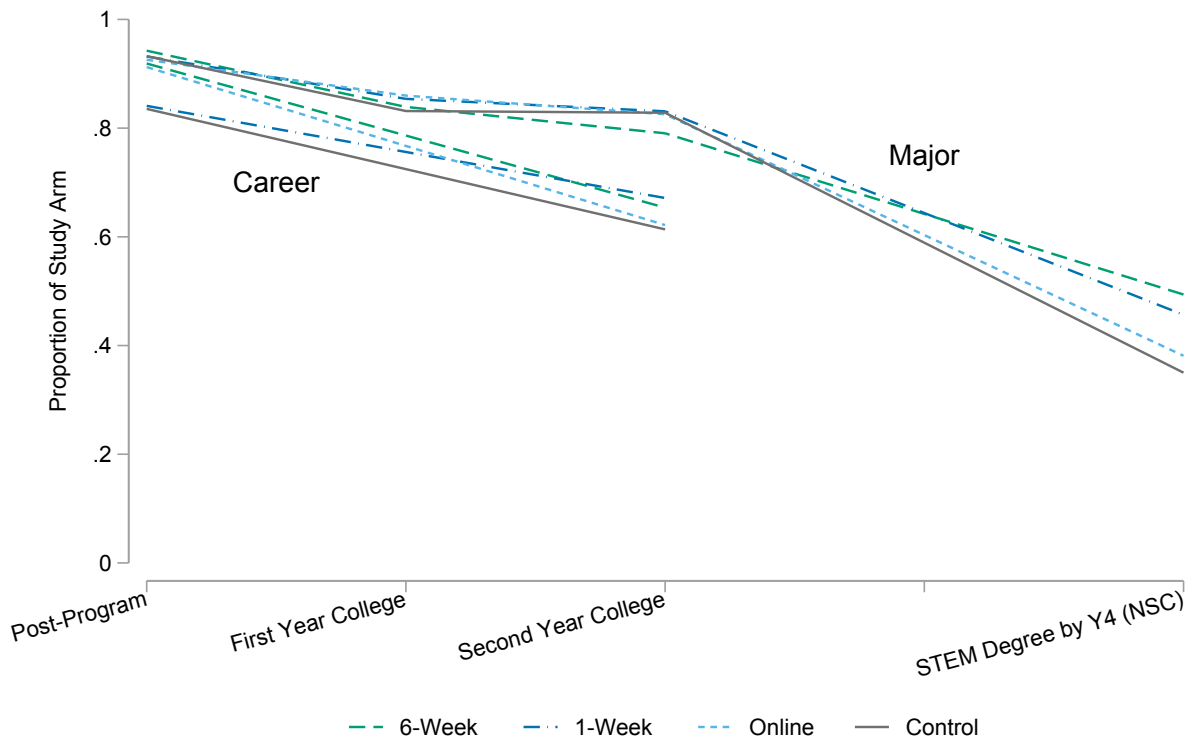
Notes: This figure displays changes in majors by program offer, with majors arrayed by “potential earnings” (Sloan et al. 2021). For detailed impact estimates on potential earnings, see Online Appendix Table B.9.

Figure 5: Actual vs. Predicted Graduation and STEM Rates



Notes: This figure compares program effects on four-year college graduation and STEM degree attainment, marked as “actual,” with “predicted” graduation and STEM degree attainment based on institution-level characteristics. The institutional-level outcomes are college-level characteristics calculated from IPEDS data in 2013. Values for community colleges and non-college-going respondents are set to 0 for both institutional-level bachelor’s four-year graduation rates and STEM degrees. For additional details, see Online Appendix Table B.16.

Figure 6: The Impact of STEM Summer Programs on 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. All responses come from surveys except for STEM degree completion which uses NSC data. For detailed impact estimates on STEM intentions, see Online Appendix Table B.17.

Table 1: Baseline Characteristics by Program Assignment

	Full Sample (1)	6-Week (2)	1-Week (3)	Online (4)	Control (5)	Strata-Adjusted Differences		
						6-Week vs. Control (6)	1-Week vs. Control (7)	Online vs. Control (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)
<i>p</i> -value						0.830	0.971	0.864
N	2084	231	308	472	1073			

Notes: This table summarizes demographic characteristics, test scores, and GPA for program applicants. Column 1 shows averages taken across the entire sample. Columns 2 through 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) were coded as not first-generation. Columns 6 through 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 are zero.

Table 2: The Impact of Assignment to STEM Summer Programs on Four-Year College Attendance

	Host Institution (1)	Any 4-Year College (2)	4-Year Excluding HI (3)	Barron's Most Competitive (4)	Barron's Most Comp. Excluding HI (5)
<u>(A) Attended in Year 1</u>					
6-Week	0.169*** (0.041)	0.038 (0.025)	-0.131*** (0.046)	0.172* (0.065)	0.003 (0.084)
1-Week	0.053 (0.038)	0.042 (0.031)	-0.011 (0.046)	0.136* (0.060)	0.083 (0.072)
Online	0.038+ (0.020)	0.020 (0.015)	-0.018 (0.030)	0.095* (0.035)	0.057 (0.049)
Control Mean	0.080	0.867	0.786	0.494	0.414
<u>(B) Attended in Year 2</u>					
6-Week	0.171*** (0.038)	0.006 (0.035)	-0.166*** (0.055)	0.126* (0.059)	-0.045 (0.079)
1-Week	0.058 (0.038)	0.033 (0.034)	-0.025 (0.050)	0.118* (0.047)	0.060 (0.062)
Online	0.043+ (0.022)	0.012 (0.023)	-0.031 (0.038)	0.097*** (0.027)	0.054 (0.043)
Control Mean	0.076	0.845	0.768	0.492	0.415
<u>(C) Attended in Year 3</u>					
6-Week	0.161*** (0.038)	-0.005 (0.035)	-0.166*** (0.060)	0.123* (0.058)	-0.038 (0.080)
1-Week	0.051 (0.038)	0.023 (0.039)	-0.029 (0.046)	0.128* (0.051)	0.077 (0.064)
Online	0.043+ (0.022)	0.005 (0.021)	-0.038 (0.036)	0.100*** (0.032)	0.057 (0.048)
Control Mean	0.078	0.815	0.737	0.469	0.391
<u>(D) Attended in Year 4</u>					
6-Week	0.178*** (0.044)	0.072+ (0.041)	-0.107+ (0.054)	0.160*** (0.053)	-0.018 (0.068)
1-Week	0.059 (0.039)	0.120* (0.044)	0.061 (0.043)	0.177*** (0.048)	0.118+ (0.060)
Online	0.047* (0.022)	0.033+ (0.017)	-0.014 (0.031)	0.112*** (0.031)	0.065 (0.046)
Control Mean	0.070	0.754	0.684	0.437	0.368

Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of the three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including indicators for GPA, standardized math score, race/ethnicity, and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014, 2015, and 2016 and passed an initial screen, who were then subject to random assignment as described in Section 4.2. The control mean is adjusted for randomization strata. Robust standard errors are in parentheses (+ $p < 0.10$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$). N = 2,084 for outcomes during the fourth year and prior, N = 1,335 for fifth year graduation, and N = 634 for sixth year graduation.

Table 3: The Impact of Assignment to STEM Summer Programs on Four-Year College Graduation

	Host Institution (1)	Any 4-Year College (2)	4-Year Excluding HI (3)	Barron's Most Competitive (4)	Barron's Most Comp. Excluding HI (5)
<u>(A) Graduated by Year 4</u>					
6-Week	0.146*** (0.035)	0.082+ (0.048)	-0.064 (0.054)	0.115+ (0.061)	-0.031 (0.069)
1-Week	0.040 (0.036)	0.080 (0.056)	0.040 (0.050)	0.099 (0.066)	0.059 (0.064)
Online	0.033+ (0.019)	0.016 (0.027)	-0.016 (0.038)	0.046 (0.039)	0.014 (0.048)
Control Mean	0.065	0.532	0.468	0.368	0.303
<u>(B) Graduated by Year 5</u>					
6-Week	0.133* (0.051)	0.122+ (0.061)	-0.010 (0.064)	0.160+ (0.083)	0.028 (0.085)
1-Week	0.022 (0.049)	0.163* (0.072)	0.141* (0.059)	0.176* (0.080)	0.154* (0.071)
Online	0.024 (0.027)	0.082+ (0.047)	0.058 (0.048)	0.117* (0.049)	0.093+ (0.050)
Control Mean	0.084	0.654	0.570	0.408	0.324
<u>(C) Graduated by Year 6</u>					
6-Week	0.204*** (0.060)	0.009 (0.062)	-0.194* (0.068)	0.046 (0.124)	-0.158 (0.105)
1-Week	-0.009 (0.053)	0.068+ (0.031)	0.077 (0.051)	0.102 (0.076)	0.111+ (0.058)
Online	0.022* (0.007)	0.098*** (0.027)	0.076* (0.032)	0.171*** (0.045)	0.149* (0.051)
Control Mean	0.086	0.736	0.651	0.431	0.346

Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of the three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including indicators for GPA, standardized math score, race/ethnicity, and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014, 2015, and 2016 and passed an initial screen, who were then subject to random assignment as described in Section 4.2. The control mean is adjusted for randomization strata. Robust standard errors are in parentheses (+ $p < 0.10$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$). $N = 2,084$ for outcomes during the fourth year and prior, $N = 1,335$ for fifth year graduation, and $N = 634$ for sixth year graduation.

Table 4: The Impact of Assignment to STEM Summer Programs on STEM and non-STEM Degrees

	Degree Within 4 Years				Degree Within 5 Years			
	Any Bachelors (1)	STEM (2)	Non-STEM (3)	Missing Major (4)	Any Bachelors (5)	STEM (6)	Non-STEM (7)	Missing Major (8)
(A) Any Four-Year Institution								
6-Week	0.082 ⁺ (0.048)	0.127* (0.057)	-0.026 (0.034)	-0.019 (0.022)	0.122 ⁺ (0.061)	0.202* (0.075)	-0.089 (0.056)	0.009 (0.045)
1-Week	0.080 (0.056)	0.092 (0.059)	-0.005 (0.027)	-0.007 (0.024)	0.163* (0.072)	0.145 ⁺ (0.083)	-0.011 (0.050)	0.029 (0.048)
Online	0.016 (0.027)	0.034 (0.026)	-0.001 (0.018)	-0.017 (0.018)	0.082 ⁺ (0.047)	0.045 (0.045)	0.014 (0.028)	0.022 (0.040)
Control Mean	0.532	0.368	0.109	0.055	0.654	0.452	0.149	0.053
(B) Host Institution								
6-Week	0.146*** (0.035)	0.145*** (0.029)	0.001 (0.019)	-	0.133* (0.051)	0.141*** (0.043)	-0.009 (0.024)	-
1-Week	0.040 (0.036)	0.056 (0.035)	-0.016 (0.014)	-	0.022 (0.049)	0.056 (0.049)	-0.034 ⁺ (0.017)	-
Online	0.033 ⁺ (0.019)	0.033 ⁺ (0.018)	-0.000 (0.006)	-	0.024 (0.027)	0.031 (0.026)	-0.007 ⁺ (0.004)	-
Control Mean	0.065	0.051	0.014	0.000	0.084	0.064	0.020	0.000
(C) Other Institutions								
6-Week	-0.064 (0.054)	-0.018 (0.055)	-0.027 (0.031)	-0.019 (0.022)	-0.010 (0.064)	0.061 (0.071)	-0.080 (0.050)	0.009 (0.045)
1-Week	0.040 (0.050)	0.036 (0.045)	0.011 (0.025)	-0.007 (0.024)	0.141* (0.059)	0.088 (0.056)	0.023 (0.046)	0.029 (0.048)
Online	-0.016 (0.038)	0.001 (0.029)	-0.001 (0.017)	-0.017 (0.018)	0.058 (0.048)	0.015 (0.041)	0.021 (0.028)	0.022 (0.040)
Control Mean	0.468	0.317	0.095	0.055	0.570	0.388	0.129	0.053

Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of the three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including indicators for GPA, standardized math score, race/ethnicity, and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014, 2015, and 2016 and passed an initial screen, who were then subject to random assignment as described in Section 4.2. The control mean is adjusted for randomization strata. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 ***p<0.001). N = 2,084 for outcomes during the fourth year and prior and N = 1,335 for fifth year graduation. Students are categorized as STEM if any of their degree majors are STEM. Degrees are categorized using CIP and, if CIP is unavailable, whether the degree is a Bachelor of Science. Students are categorized as STEM if at least one major is STEM. Students are categorized as non-STEM if none of their majors is STEM. Students who attained degrees, but have no degree major codes and did not attain Bachelors of Science, are categorized as Missing Major.

Table 5: The Impact of Assignment to STEM Summer Programs on Self-Reported College Applications and Admissions

	Any 4-Year (1)	Number of Institutions (2)	Number of Institutions Excluding HI (3)	One Institution College (4)	Barron's Most Competitive Except HI (5)	Barron's Less Competitive (6)	Technical School (7)	Technical School Except HI (8)	State Flagship (9)
(A) Applications									
6-Week	0.006 (0.009)	0.148 (1.184)	-0.179 (1.164)	-0.094*** (0.030)	0.113* (0.041)	-0.029 (0.025)	0.247*** (0.054)	0.039 (0.082)	0.026 (0.060)
1-Week	0.006 (0.010)	0.342 (0.791)	0.141 (0.766)	-0.041 (0.029)	0.062 (0.049)	-0.029 (0.020)	0.181*** (0.041)	0.101 (0.066)	0.027 (0.060)
Online	-0.003 (0.007)	0.822+ (0.406)	0.571 (0.397)	-0.032* (0.013)	0.043+ (0.023)	-0.015 (0.012)	0.182*** (0.029)	0.080* (0.038)	0.026 (0.036)
Control Mean	0.992	8.480	8.078	0.077	0.841	0.059	0.600	0.458	0.473
Observations	1418	1354	1354	1354	1407	1407	1407	1407	1407
(B) Admissions									
6-Week	0.006 (0.009)	0.002 (0.792)	-0.154 (0.780)	-0.069 (0.043)	0.103+ (0.054)	-0.018 (0.025)	0.145* (0.066)	0.011 (0.072)	0.063 (0.065)
1-Week	0.006 (0.010)	0.322 (0.702)	0.267 (0.688)	-0.012 (0.031)	0.096 (0.060)	-0.020 (0.019)	0.106+ (0.062)	0.084 (0.062)	0.036 (0.067)
Online	-0.003 (0.007)	0.276 (0.281)	0.209 (0.268)	-0.015 (0.018)	0.081* (0.030)	-0.014 (0.013)	0.093* (0.044)	0.035 (0.043)	0.026 (0.044)
Control Mean	0.992	5.382	5.226	0.124	0.688	0.055	0.475	0.398	0.418
Observations	1418	1400	1400	1400	1402	1402	1402	1402	1402

Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of the three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including indicators for GPA, standardized math score, race/ethnicity, and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014, 2015, and 2016 and passed an initial screen, who were then subject to random assignment as described in Section 4.2. The control mean is adjusted for randomization strata. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 ***p<0.001). Data are from surveys conducted at the end of the senior year of high school. The any application and any admission outcomes were created from yes or no questions and available for all survey respondents. The number of applications and admissions outcomes in Columns 2 through 4 were calculated for respondents who provided their full list of applications and admissions, respectively. The applications and admissions to types of institutions outcomes in Columns 5 through 9 are populated for respondents who provided at least a partial list of applications and admissions, respectively.

Table 6: The Impact of Assignment to STEM Summer Programs on AP/IB Courses and Skills

Panel A: Plans to Take AP or IB Courses						
	Any Course (1)	Number of courses (2)	STEM (3)	Science (4)	Computer Science (5)	Math (6)
6-Week	0.063 (0.043)	0.140 (0.322)	0.087+ (0.045)	0.026 (0.071)	0.097+ (0.050)	-0.050 (0.071)
1-Week	0.112* (0.041)	0.713+ (0.351)	0.149*** (0.045)	0.058 (0.067)	0.078 (0.057)	0.060 (0.067)
Online	0.027+ (0.015)	0.284 (0.187)	0.050* (0.021)	0.008 (0.042)	0.073 (0.047)	0.025 (0.042)
Control Mean	0.891	4.281	0.842	0.736	0.104	0.734
Observations	1057	1057	1057	1057	1057	1057

Panel B: Skills and Confidence						
	Calculus Question (7)	Life Skills Index (8)	Study Skills Index (9)	Confidence Index (10)	Likes Intellectual Activities (11)	Attention Span (12)
6-Week	0.029 (0.066)	0.378* (0.141)	0.441*** (0.125)	0.028 (0.165)	0.171 (0.139)	-0.208 (0.152)
1-Week	0.036 (0.051)	0.267+ (0.143)	0.369*** (0.088)	0.134 (0.143)	0.065 (0.117)	-0.341* (0.128)
Online	0.039 (0.038)	0.116 (0.100)	0.268*** (0.058)	-0.017 (0.105)	0.096 (0.077)	-0.224* (0.090)
Control Mean	0.213	-0.029	-0.034	-0.044	-0.029	0.085
Observations	1018	1388	1380	1031	1033	1379

Notes: Each coefficient labeled by program is the estimate of the impact of assignment to one of the three STEM summer programs on the outcome indicated in the heading. All regressions control for randomization strata and a vector of characteristics including indicators for GPA, standardized math score, race/ethnicity, and free and reduced-price lunch status. The sample includes STEM summer program applicants who applied in 2014, 2015, and 2016 and passed an initial screen, who were then subject to random assignment as described in Section 4.2. The control mean is adjusted for randomization strata. Robust standard errors are in parentheses (+ p<0.10 * p<0.05 ** p<0.01 ***p<0.001). Data are from surveys conducted post program, in the fall of the senior year of high school. Data are only available for the 2015 and 2016 cohorts for AP/IB and confidence and likes intellectual activities outcomes.