Early Exposure to Technology and College Major Choice*

Edoardo Bella ⓡ Garrett Anstreicher ⓡ Joanna Venator†

June 29, 2026

Abstract

We leverage the national rollout of FIRST robotics teams across high schools and document that attending a high school with a robotics club substantially increases the likelihood that a student will go on to major in computer science or engineering in college. These effects vary significantly by gender: while these programs increase the likelihood that male students go on to major in engineering/computer science (+3.1 pp), effects for female students are substantially smaller and are at times negative, with worse effects for teams with lower gender ratios. We confirm these findings in restricted-use Census survey data and find that positive impacts extend past college, with robotics teams increasing the likelihood that men work in STEM-related occupations. These results suggest that exposure to STEM-related high school extracurricular activities may increase interest in engineering and computer science but can also widen existing gender gaps in take-up of these fields.

Keywords: College Major Choice, Robotics, High School Curricula, Occupation Choice

JEL: I23, J24, J16

*We thank conference participants at SOLE for helpful feedback and thank Abigail Cheng, Mai Dasgupta, and Molly Wysocki for excellent research assistance. The authors have no sources of research support/any financial relationships for this paper to disclose. Corresponding Author: Joanna Venator, 140 Commonwealth Ave, Chestnut Hill, MA 02467; venator@bc.edu. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 3098. (CBDRB-FY26-P3098-R12992).

†Affiliations: Anstreicher: University of Iowa; Bella and Venator: Boston College

1 Introduction

A recent strand of literature in the economics of education has highlighted the importance of college major choice as a determinant in later-life labor market outcomes — indeed, some work has suggested that this margin of decision may be as impactful as the choice to attend or to complete college at all [Altonji et al., 2012, Andrews et al., 2024]. Nonetheless, estimates of the elasticity of major choice with respect to future earnings are relatively small, and even accounting for differences in beliefs about post-college earnings and abilities, much of the gender gap in take-up of high-paying majors such as engineering and computer science are unexplained [Patnaik et al., 2021]. A better understanding of the factors that encourage individuals to pursue higher-paying fields is thus of clear interest to the current policy dialogue surrounding improving career pathways both overall and for underrepresented groups.

In this paper, we identify a new driver of students' interest in STEM majors — high school extracurricular activities — and provide the first causal estimates of their importance. Because engineering has a particularly rigid course structure relative to other STEM fields and often requires students to declare interest as freshmen, early exposure to coding and robotics through extracurriculars may make this major more accessible to students who otherwise would not pursue this major. We leverage the roll-out of the earliest high school robotics competition, For Inspiration and Recognition of Science and Technology (FIRST) robotics, across high schools to test how attending a high school with a robotics team impacts students likelihood of both majoring in engineering or computer science and working in STEM in the future.

The FIRST robotics competition began in 1992 and has since grown to include over 3,000 teams from more than 30 countries. We build a database of all robotics teams from 1992 to present and merge it with college transcript data from College and Beyond II, an administrative database of college student record and transcript data for over 1.3 million students from 19 public colleges across seven university systems. Using a bias-adjusted two-way fixed effects design, we show that students from schools with robotics teams are 1.1 pp more likely to major in computer science, with larger impacts for students at schools with well-established teams. In the full sample, there are null effects on take-up of engineering majors and STEM majors more generally.

These null effects hide substantial heterogeneity by student's demographic characteristics. For male students, there is a large and statistically significant impact of attending a high school with a robotics team on take-up of engineering and computer science majors (3.1 pp), and computer science majors alone (1.9 pp). In contrast, female students have a null or even negative impact of robotics teams on their likelihood of choosing these majors. To test how demographic make-up of the team impacts student outcomes, we survey robotics team on their gender make-up and use the relative participation rate of women by team to show that majority male teams are particularly discouraging for female students. Teams that are less than 20% female induce women to be 2.7 pp less likely to major in engineering and computer science whereas teams that are more than 20% female increase women's engineering and computer science major take-up by 2.3 pp. In contrast, men's take-up of engineering and computer science majors is unaffected by gender make-up of the team.

Finally, we investigate whether this increased major take-up carries into the labor force by merging our database of robotics teams with restricted-use Census Survey data that contains information on college major, labor market outcomes, and county of residence in 2000. We obtain quantitatively and qualitatively similar patterns of heterogeneity in the impacts of exposure to robotics teams on college major choice and find that, mostly for men, exposure to robotics teams is associated with an increased likelihood of working in STEM-related occupations. However, these impacts do not appear to translate into higher hourly wage earnings, suggesting that these college major and occupation substitution patterns may be primarily occurring among individuals who would have performed well in the labor market regardless. Taken together, these results suggest that exposure to STEM-related high school extracurriculars can increase interest in technology-related fields but can also unintentionally widen the already-existing gender gap in interest in them.

These results contribute to several strands of economics and education literature. First, our paper demonstrates a new channel through which students' gain field-specific preferences, speaking to the literature on determinants of college major choice.[1] Within this literature, economists have documented effects of ability [Arcidiacono, 2004] as well as expected earnings on what fields students choose to study [Wiswall and Zafar, 2015a,b, Conlon, 2021, Conlon and Patel, 2026], while other research has documented that major choice is sensitive to overall macroeconomic conditions [Blom et al., 2021]. However, there remains a substantial unexplained portion of preferences for college major that forms prior to college matriculation, which has given rise to attempts to understand the pre-college factors that influence college major preferences. While some results in this endeavor suggest that high school curricula do not appear to be significant drivers of future STEM career decisions [Darolia et al., 2020],

[1]See Patnaik et al. [2021] for a review of this literature

other entries point to parental background [van der Vleuten et al., 2018] and the gender composition of high school teachers [Bottia et al., 2015] playing a meaningful role. In our focus on extracurriculars, we contribute to the ongoing effort to unpack this "black box."

Pre-college factors may be particularly important for majors such as engineering which often require students to apply for entry into the engineering program when applying for college and have strict course requirements that make on-time graduation difficult for those who switch into the major mid-college. This creates a form of path dependence: if a marginal student receives training or information about engineering prior to high school, it may push them onto a path towards occupations with long run earnings gains whereas students who aren't exposed to engineering in high school may miss the window where they can select into this field. In this sense, our work has similarities to the literature which document how timing of specialization into field – pre-college versus during college – impacts long-run outcomes [Kemple and Willner, 2008, Dougherty, 2018, Brunner et al., 2023, Bordon and Fu, 2015, Fricke et al., 2018, Malamud, 2010, Dahl et al., 2023, Bruhn et al., 2025].

The heterogeneity in effects over gender that we uncover also contribute to the large literature that has attempted to better understand and address large gender-based disparities in STEM interest, which itself plays a meaningful role in the gender earnings gap [Zafar, 2013]. Past work in this area has focused on factors such as the demographics of faculty and teaching assistants [Price, 2010, Carrell et al., 2010, Griffith, 2014, Griffith and Main, 2021], adult role models [Breda et al., 2023], and intergenerational effects [Humlum et al., 2019, Philipp, 2023]. Notably, this literature has highlighted choices made in high school specifically as important drivers of this gap later in the lifecycle [Brenøe and Zölitz, 2020, Card and Payne, 2021, Gurantz, 2021, Granato, 2023, Cohodes et al., 2026], and similar gendered patterns of heterogeneity have been observed in other policy contexts, such as Career and Technical Education (CTE) programs [Brunner et al., 2023]. Engineering and computer sciences are fields within STEM in which gender gaps are particularly persistent, and our work provides new evidence on the role that high school extracurriculars may play in driving this gap.

Lastly, this paper is the first to evaluate the causal impacts of the roll-out of FIRST across high schools on college major choice using administrative college transcript data, rather than small-scale survey based measures. Several studies have assessed the impact of high-school robotics competitions (or comparable initiatives) on participant outcomes such as attitudes toward STEM or information-technology-related fields [Welch and Huffman, 2011, Mason et al., 2011, Jackson et al., 2019][2]. Notably, Meschede et al. [2024b] aim to study the long-run impacts of participation in FIRST competitions by comparing the outcomes of approximately 800 FIRST participants with 450 comparison students. While the authors attempt to match participants with comparisons along observable lines, they are unable to control for selection on unobservables into FIRST participation in the first place. We complement and expand upon this effort through bringing frontier econometric methods to answer the research question and situate our findings relative to theirs later in the paper.

The rest of the paper proceeds as follows. In Section 2, we describe the history of FIRST robotics competitions and our data on their expansion. Section 3 describes our empirical approach for estimating the causal impact of robotics competitions on long-term STEM interest, and Section 4 presents our empirical results. Sections 5 and 6 discuss our findings in the context of the broader literature and conclude.

2 Data and Policy Setting

2.1 FIRST Robotics

FIRST Robotics is an international non-profit founded in 1989 with the goal of encouraging students to enter engineering and technology fields through programs which introduce robotics to high school and middle school students. Their primary program, the FIRST Robotics Competition (FRC), is a program in which teams of high school students build robots to compete in annual challenges that require the robots to complete tasks such as scoring goals or completing obstacle courses. Students build the physical chassis for a robot, design and maintain the electrical and pneumatic systems, and program the code used to run the autonomous systems in the robot. Most teams are extracurricular after-school clubs rather than classes and are guided by both teachers within the high school and industry mentors.

Using publicly available data scraped from the FIRST website's competition archives, we create a panel of team-by-year information on participation in FIRST. While the original FRC took place in 1992 with 28 teams, it has rapidly expanded in the last thirty years, with the most rapid expansion occurring in the mid 2000s. In 2000, there were 372 teams

[2]See also Anwar et al. [2019] for a broader review of this literature. Beyond robotics initiatives specifically, other papers have also assessed broader high school environments in influencing STEM career choice, including AP class offerings [Long et al., 2012, Conger et al., 2021] and high-school math curricula [Aughinbaugh, 2012]

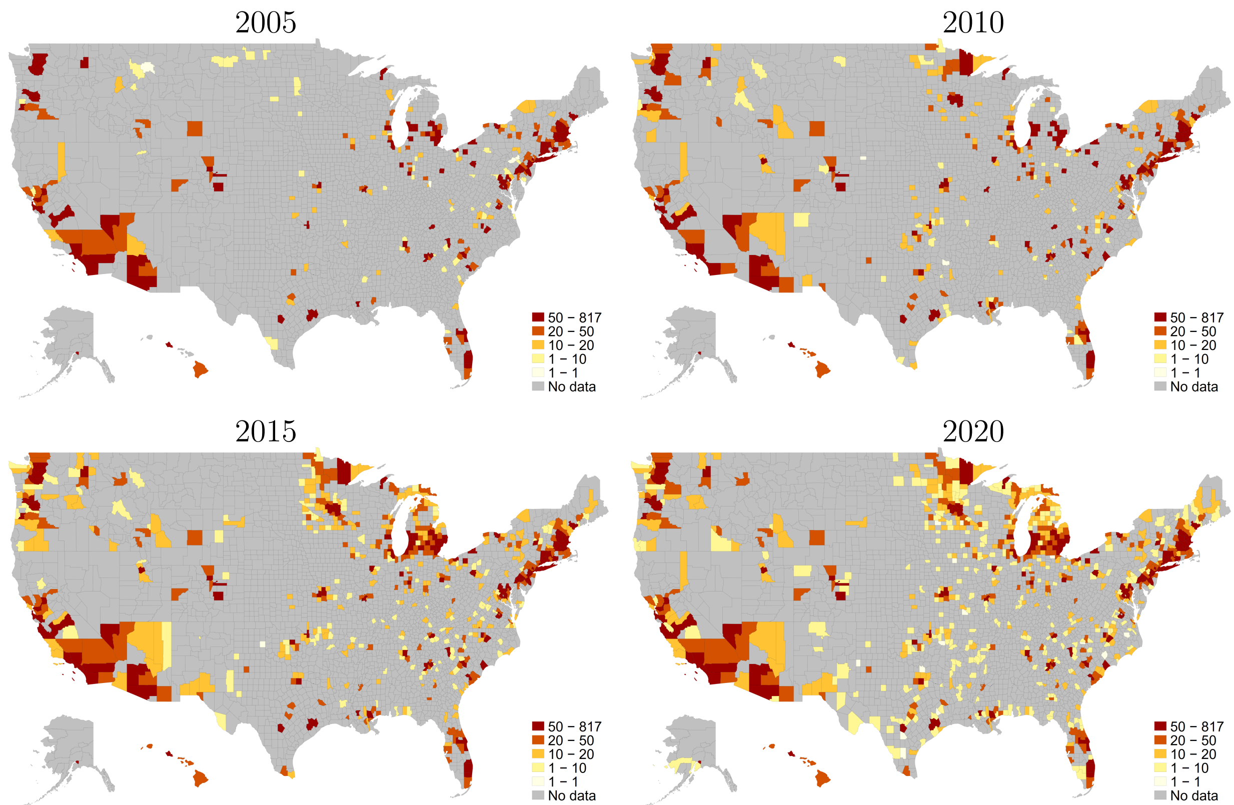


Figure 1: Number of robotics teams by county, 2005-2020

competing in FRC. By 2010, there were 1808 teams including over 45,000 high school students competing in over 53 different competitions. As of 2023, there were 3,304 teams involving over 83,000 high school students in 167 different competitions.

Figure 1 shows a heat map of the number of teams per county in 2005, 2010, 2015, and 2020. As shown in the upper left graph, early adoption was concentrated primarily in the Northeastern corridor (Boston-NY-Philadelphia-DC) and in technology/engineering hubs such as Silicon Valley, Seattle, and Detroit. By 2015, the program had spread beyond technological hubs and is fairly widespread over most regions excluding the less populous states in the Great Plains/ Mountain region (e.g., the Dakotas, Montana, Wyoming)

For each FRC team, we are able to observe the school name(s), city, state, and year that the team originally competed (rookie year henceforth). We create a panel of schools which had a FRC team from 1992 through 2021. Because the data does not include unique identifiers for school name, we use a fuzzy matching algorithm to match school names reported in the raw data set to school names in the database of CEEB codes provided by the College Board. We first drop any teams which do not report a rookie year or report that they are an international team. We standardize common words (e.g., high school and HS), remove punctuation and stop words, and then match within state on school name to the CEEB database using Stata18 command matchit. In cases where a team has multiple organizations attached to it, we separate the organizations into separate strings and match each organization individually. Out of 5,803 teams in our scraped database, we are able to get a match to at least one school name for 3,631 or 62.5% of teams. In total, we have 4,116 high schools which ever were associated with an FRC team.[3]

2.2 College and Beyond II

Our first source of outcome data comes from the College and Beyond II (CBII), a dataset administered by the Inter-university Consortium for Political and Social Research (ICPSR) including college student record and transcript data for over 1.3 million students from 19 public colleges across seven university systems including the City University of New York system, Georgia College and State University system, Indiana University-Bloomington, Truman State University, University of California-Irvine, University of Houston, and University of Michigan. The data covers all students who attended these universities from 2000 through 2020 and provides information on courses taken, grades, college completion, major and minor choice, demographics (sex, race, citizenship), and pre-college characteristics including pre-college test scores, family income, permanent address zip code, and high school code.

We merge this data with our school-year panel of robotics team data using year of high school entry, measured as four years prior to entry to college, and high school CEEB code when available.[4] In the final data set, we know for each student if they studied in a high school that ever participated in the program, if they entered the school after the program was already in place, and for how many years they were exposed to the competition. We define a student as treated based on the treatment status of their high school in their freshman year (i.e., treated for all of high school).

Our primary outcome of interest is student's college major at time of graduation. As we would expect that exposure to robotics extracurriculars increases interest and/or skills in STEM fields, such as computer science or engineering, we create three measures of STEM majors: an indicator for majoring in any STEM field, defined as CIP codes 1, 11, 14, 15, 26, 27, 40, 41, or 51, an indicator for majoring in engineering or computer science, defined as

[3]There is a higher number of high schools than teams because a team can be associated with multiple high schools.

[4]CEEB is provided in only three out of the seven university systems.

CIP codes 11, 14, or 15, and an indicator for majoring in computer science alone, defined as CIP code 11.[5] In supporting analyses, we look at the impacts on majors we would expect students to be drawn out of, such as humanities majors (CIP codes 5, 16, 22, 23, 24, 25, 38, 39, 50, and 54) and social science majors (CIP codes 4, 9, 10, 91,42, 44, 45, and 52). We also are able to look at proportion of courses associated with each major grouping and grades in these courses.

2.3 American Community Survey

While our data from CBII contains many useful outcome measures, it also comes with important limitations. The fact that all the data in CBII are for students *in college* means we cannot say whether any major substitution patterns we observe carry into the labor force, nor can we estimate extensive-margin effects (that is, students selecting into STEM majors who otherwise may not have attended college at all). Moreover, the fact that CBII comes from only seven university systems may give rise to concerns regarding representativeness and the possibility that our results are dominated by a single large institution.

To overcome these limitations and to provide an assessment of the robustness of our results across datasets, we leverage restricted-use data from the American Community Survey (ACS), administered by the Census Bureau. The ACS is an annual survey administered to one percent of the U.S. population that contains a detailed set of questions regarding geography, education, and labor market outcomes. A key limitation of the public-use data, however, is that the only piece of historical geographic information for a respondent is their state or country of birth, which would prevent us from determining their FIRST treatment status with any degree of confidence. To overcome this limitation, we link individuals to their Protected Identification Keys (PIKs; essentially scrambled Social Security Numbers) from the Social Security Administration NUMIDENT file and then link individuals *to themselves* in the 2000 Decennial Census, thereby obtaining detailed information on their location in 2000.

We use the 2009-2024 waves of the ACS, as the ACS began to record primary and secondary fields of study for college graduates in 2009. We focus on the birth cohorts of 1982 to 1999, such that respondents would be at most 18 in 2000 and still attached to their parents, thereby preempting the most migratory point of the lifecycle. We also limit our sample to individuals aged at least 25 when surveyed to increase our confidence in

[5]We do not look at engineering majors alone because a subset of the colleges in CBII do not offer engineering majors.

respondents having completed their undergraduate schooling.

We link individuals to their FIRST treatment at the year-2000 county level, defining treatment based on whether this county had at least one treated high school during their freshman year as above.[6] As above, we classify majors as STEM as well as specifically in engineering or computer science,[7] and we classify individuals as working in STEM-related fields based on their occupation code and Census occupation crosswalks.[8]. Own-income and occupation are both measured at time of survey (i.e., between age 25 and 42 depending on survey year and birth cohort). Table B.I presents summary statistics of this sample, both overall and then separately by year-2000 parent income for individuals surveyed specifically in the 2000 Long-Form. While women in the sample attain more schooling on average overall, rates of STEM degree attainment are roughly equal across sexes, and men are more likely to work in STEM fields and obtain BAs in engineering and computer science specifically.

3 Empirical Methods

We estimate the impacts of high school robotics teams on college outcomes using a staggered two-way fixed effects specification. Because the roll-out of FRT across school districts is staggered across time, a standard two-way fixed effect specification may be subject to bias if there are heterogeneous treatment effects by event-time. We therefore use the imputation-based estimator from Borusyak et al. [2024] which fits the high school and graduation cohort fixed effects using untreated observations, uses these fixed effects to impute a treatment effect for each period, and then takes a weighted sum of these treatment effects to aggregate into a single treatment effect. We use both a static (eqn. 1) and a dynamic event-study framework (eqn. 2)

$$Y_{ist} = \beta \mathrm{FRT}_{s(i),t} + \gamma X_i + \theta_{s(i)} + \delta_t + \alpha_{g(i)} + \varepsilon_{ist} \tag{1}$$

$$Y_{ist} = \sum_{\tau=-16}^{+16} \beta_\tau \mathbb{1}(t = \text{FRT rookie year} + \tau) + \gamma X_i + \theta_s(i) + \delta_t + \alpha_{g(i)} + \varepsilon_{ist} \tag{2}$$

[6] A limitation of the ACS is that we cannot observe the high school that an individual attended, meaning that our treatment is measured with noise. For this reason, we prefer the CBII as our main specification for major outcomes and see the ACS results primarily as a validation exercise.

[7] We do not focus on computer science specifically in these data, as respondent counts for this specific major over event time become quite small, which raises disclosure avoidance issues.

[8] See https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html

Y_{ist} is a dummy for whether student i who attended high school s and was a high school freshman in year t majored in a given major category. $FRT_{s(i),t}$ is our treatment variable equal to one if school s had a team in year t. In the dynamic setting, we include indicators for leads and lags around the school's rookie year (i.e., $t = 0$) with the year prior to the school starting a team as the omitted category. X_i are student-level controls including gender, indicators for race, indicators for US citizenship, indicators for parental income above $100 thousand dollars, and indicators for at least one parent having a bachelor's degree. We also include fixed effects for student's high school ($\theta_{s(i)}$), student's graduation cohort (δ_t), and student's home state ($\alpha_{g(i)}$). Standard errors are clustered at the high school level. Analyses using the ACS data proceed equivalently, except that the relevant geographic fixed effects and clustering level are at the county level, and fixed effects for home state and parental covariates are omitted. In the dynamic setting, a county's "rookie year" is the year that the county first had a team.

To interpret the β as causal estimates of the impact of robotics teams on college major choice, it must be the case that students from untreated schools (counties) follow similar trends in the pre-period in term of outcomes and that the timing of their high school (county) starting a FRT is uncorrelated with unobserved time-varying determinants of college outcomes. We might be concerned that high schools that perceive greater demand for technology skills, either from the surrounding labor market or from student preferences, are more likely to start a team. For example, earlier adopters of FIRST robotics tended to be in areas with industry partners willing to sponsor schools' teams; Michigan has higher rates of FIRST participation in part due to buy-in from the auto industry. While time or location invariant differences will account for cross-sectional time invariant differences or time-invariant high school differences, we also explore whether observable characteristics about students seem to vary with timing of FRT acquisition.

Table 1 reports the mean values of demographic characteristics for students in treated schools (col. 1), students in control schools (col. 2), and the results of a regression to test balance in demographic characteristics and high school test scores for students in treated schools relative to those untreated (col 3). We run our primary specification, a two-way fixed effects model adjusted using the method from Borusyak et al. [2024], with the demographic characteristic as the outcome. While there are statistically significant differences between treated and untreated students, they are mostly small in actual value relative to baseline. The exception is the proportion of Asian students and the proportion of students whose parents have a BA; treated students are 2.2 pp more likely to be Asian and they are 6.8 pp more likely to have parents with a college degree.

Table 1: Summary Statistics, by Treatment

Variable	Mean, Treated	Mean, Control	Difference
Female	0.532	0.543	-0.014*
White	0.589	0.620	-0.010
Asian	0.177	0.123	0.022***
Hispanic	0.094	0.121	-0.010*
Black	0.053	0.061	-0.008
Citizen	0.954	0.951	0.009*
Income > 100 K	0.546	0.415	0.003
Income Missing	0.282	0.405	-0.065***
Parent BA	0.692	0.491	0.068***

Table Notes: This table reports the mean values for students at treated schools and controls schools. Treated means that the student attended a school that had a team in the previous 4 year before their college first term. Column 3 reports the coefficient from a two-way fixed effects Borusyak et al. [2024] adjusted model with the demographic type as the outcome, high school and cohort FEs, and controls for other demographics. Standard errors clustered at the high school level are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Next, we look at aggregate economic conditions in the labor markets that teams are located in. For this we use school-level characteristics from 2009-2019 sourced from the Stanford Education Data Archive 5.0 [Reardon et al., 2024] cross-walked to high school CEEB code using crosswalks from [Lyles et al., 2020]. We regress school outcomes (total enrollment, percent white, percent black, percent free or reduced price lunch, rural status, magnet, charter) on an indicator for if the school had a robotics team along with high school, year, and state fixed effects. We see that schools with robotics teams are slightly larger (+1.4% relative to baseline mean), have fewer White students (-0.7%) and more Black students (+3.0%), have more students on free or reduced price lunch (-0.9%), and are more rural (-0.5%). While these differences are statistically significant, they are all small relative to the baseline means.

Taken together, these results suggest that students and schools exposed to robotics teams are not fully random, but the differences are small in aggregate. For this reason, we control for demographics in all specifications, as well as looking at whether there are differential impacts of treatment by demographic types. Our identification strategy thus relies on conditional parallel trends assumptions; that is, controlling for demographic characteristics, the path that college major take-up would take in the absence of robotics entry is similar to that

Table 2: Balance of School Level Characteristics

	(1) Enrollment	(2) % White	(3) % Black	(4) % FRPL	(5) Rural	(6) Charter	(7) Magnet
Treated	12.79**	-0.005***	0.004***	-0.004**	-0.022***	0.001	-0.006
	(5.401)	(0.001)	(0.001)	(0.002)	(0.007)	(0.001)	(0.004)
Mean	866.77	0.638	0.130	0.460	0.420	0.010	0.045
N	151595	151595	151595	151595	151595	151595	151595
R^2	0.979	0.989	0.991	0.931	0.893	0.927	0.813

Table Notes: This table reports a regression of school-level outcome variables sourced from Stanford Education Data Archive on an indicator for being treated, along with high school FE, year FE, and state FE. Treated means that the school that had a team in the year that the statistic is calculated. Standard errors clustered at the high school level are in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

of the control group. We follow Borusyak et al. [2024]'s suggested method for constructing pre-trend estimates in all dynamic specifications to assess the validity of this assumption directly.

4 Results

4.1 Impacts on College Major

Our first set of results looks at whether exposure to high school robotics changes students' major choice. Table 3 reports the results of these regressions for the full sample and separately by demographic group.[9] There are statistically insignificant effects on overall STEM major for all groups, but noisy coefficients with opposite signs across demographic groups, suggesting heterogeneity in effects by gender in particular. When we drill down into STEM majors directly related to robotics, we see that men increase take-up of CS/engineering majors by 3.1 pp (+11% off a baseline of 26.9pp), which is primarily driven by an increase in CS majors of 1.9 pp (+29% off a baseline of 6.5pp). For women, there is no significant effect of attending a high school with a robotics team on take-up of engineering or CS majors. In a t-test of equality of coefficients, we can reject the null that the coefficients are equal across genders for take-up of CS/Engineering majors at the $p < 0.05$ level.

[9] In Appendix Section A, we report results from a standard two-way FE model not adjusted for staggered effects for comparison. The results are similar, though the differences by demographic type are amplified in the standard two-way FE model.

Table 3: Impacts of Robotics Team Entry on Major Take-up

	(1) STEM Major	(2) CS or Engineering	(3) CS
Full Sample	-0.0250	0.0129	0.0114**
	(0.0250)	(0.00997)	(0.00516)
N	97,221	97,221	97,221
Baseline Mean:	0.344	0.165	0.037
Men	0.0137	0.0306*	0.0190**
	(0.0155)	(0.0158)	(0.00829)
N	44,579	44,579	44,579
Baseline Mean:	0.420	0.269	0.065
Women	-0.0314	-0.00806	0.00523
	(0.0263)	(0.00867)	(0.00476)
N	51,965	51,965	51,965
Baseline Mean:	0.279	0.076	0.013
Parent BA	-0.00245	0.00781	0.0136***
	(0.00839)	(0.00953)	(0.00496)
N	82,399	82,399	82,399
Baseline Mean:	0.320	0.163	0.0516
Parent Non-BA	0.00431	0.00479	0.00850*
	(0.00882)	(0.00810)	(0.00464)
N	78,944	78,944	78,944
Baseline Mean:	0.279	0.106	0.0307
Family Income > 100 K	-0.0132	-0.00898	0.00674
	(0.0155)	(0.0148)	(0.00672)
N	32,648	32,648	32,648
Baseline Mean:	0.322	0.171	0.0558
Family Income < 100 K	0.00680	0.0144	0.0149***
	(0.0120)	(0.0101)	(0.00576)
N	42,765	42,765	42,765
Baseline Mean:	0.292	0.135	0.0402

Notes. This table reports regression coefficients from regressions of indicators for major (STEM: CIP 1, 11, 14,15, 26, 27, 40, 41,51; CS or Engineering:11, 14, 15; CS: 11) on the treatment (school having a robotics team) in a bias adjusted regression using the method from Boruysak et al. implemented using Stata program did-imputation. Baseline mean is the average value of the outcome for untreated students. This includes controls for demographics, high school and graduation cohort FE and clusters standard errors at the school level;* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

When we look separately by socioeconomic background, we see a positive, significant effect of treatment on majoring in CS across parental education backgrounds, with a larger coefficient for students whose parents have a BA. Conversely, while the coefficient is positive for students whose parents make more or less than $100,000, it is only statistically significant for students whose parents make less than $100,000. These students are 1.5 pp more likely to major in CS, or 37.5% relative to the baseline rate of 4.0 pp. Though we cannot reject the null that the levels of effects are equal across parental education or parental income, the lower baseline rate of take-up for students whose parents have less education or lower income means that the percent change in major choice is larger for these groups.

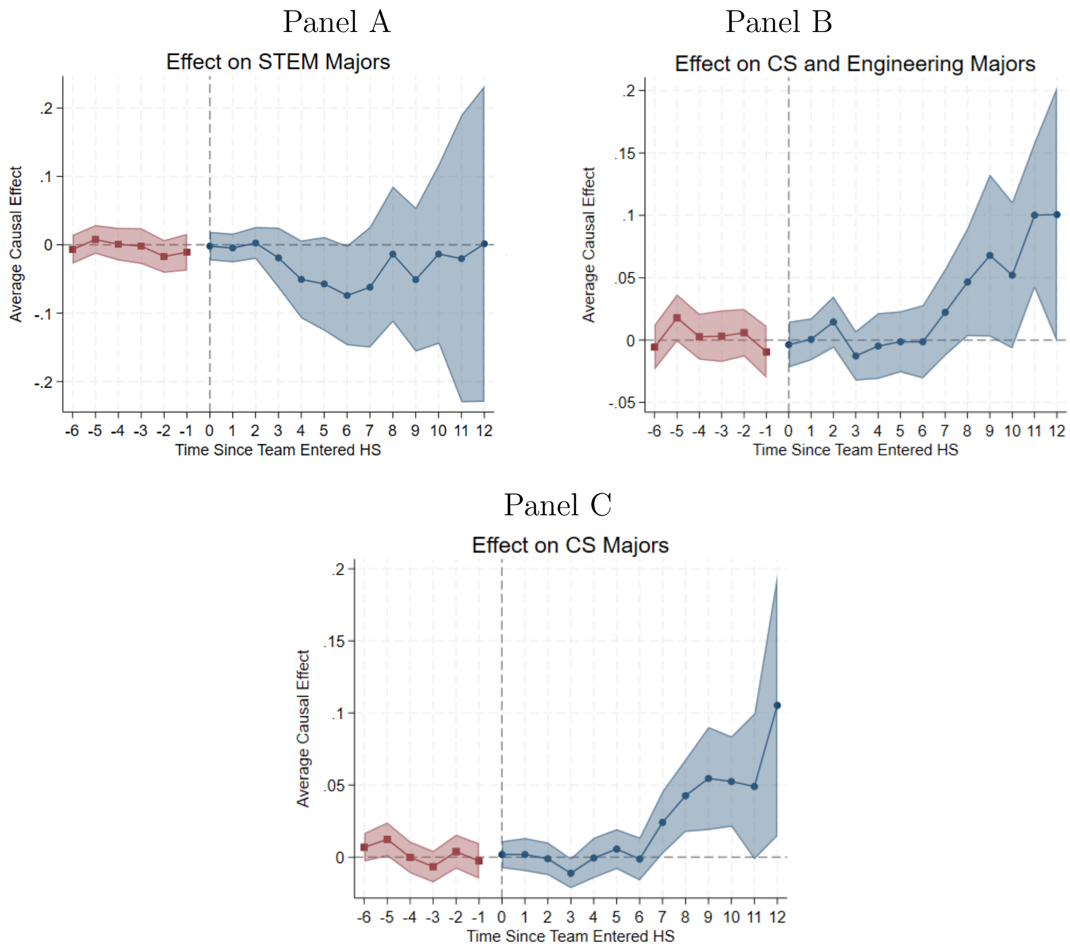
We next turn to event study specifications using the dynamic imputation-based estimator from Borusyak et al. [2024] to test for parallel pre-trends and explore how timing of a team's entry into a school impacts students' major choices. For this, we focus on the full sample and the differences by gender.[10] Given that the robotics teams require funding and institutional knowledge which demand some lead time to build up, we may expect that the earlier years of a robotics team are less effective at encouraging students to go into related fields, especially if the early years of operation of a team are less successful and enjoyable for the participants. Figure 2 show the results of regressing indicators for majoring in STEM (panel A), engineering/computer science (panel B), and computer science specifically(panel C) on dummies for timing of robotics team entry relative to a student's freshman year of high school, as specified in equation 2.

For all three outcomes, we see parallel trends in the pre-period. We see null effects in the overall sample for STEM majors. For CS and Engineering majors and CS majors alone, we see an increase in likelihood of students choosing technology majors after a delay of about 6 years, suggesting that dynamic effects are a salient feature of our results. Students whose school has had a FRT for eight years when they enter high school are about 5 pp more likely to choose an engineering or CS major, whereas students whose school has had a team for 12 years are approximately 10 pp more likely to choose an engineering or CS major.

Figure 3 shows the same specifications separately by gender. While the disaggregated time periods are estimated noisily, the patterns are consistent with the results found in the static specification: going to a school with a robotics team increase male students' take-up of STEM majors, whereas it has a null or even negative effect on women's take-up. We do see a statistically significant increase in CS majors for both men and women, but the effect

[10]We focus on gender rather than parental income or education because there was a significant difference in post-period outcomes in the aggregate for gender, but not by SES.

Figure 2: Robotics Teams Increase Likelihood of Majoring in STEM

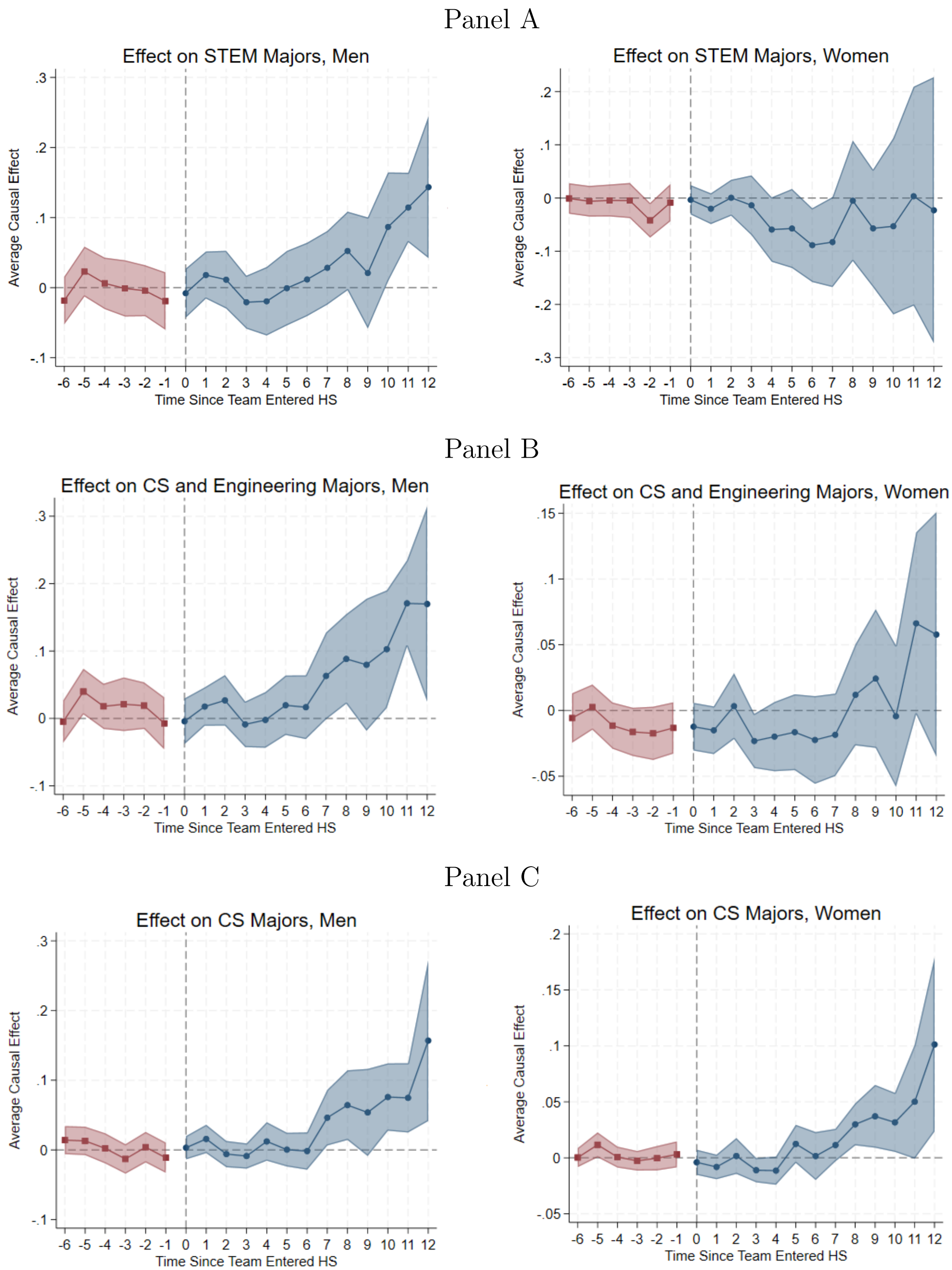


Notes: This figure reports coefficients based on equation 2, our event study regression, estimated using imputation adjusted methods from Borusyak et al. [2024], for the outcomes majoring in STEM (panel A), engineering/computer science (panel B), or computer science (panel C). All regressions include fixed effects for high school, state, and student's graduation cohort and controls for student gender, race, parental income, parental education, and student citizenship. 95% confidence intervals included; Standard errors clustered at the high school level.

sizes for women are approximately half the size of the increases for men.

In all specifications, it takes 6-7 years for a robotics team to impact students' outcomes. This demonstrates the importance of accounting for heterogeneity of treatment effects over event time. We attribute the delayed effect to schools needing time to develop the team — it takes time for the team to build membership, and it is rare for teams to be successful at the robotics competitions in their early years of participation. To provide further evidence for this hypothesis, we observe teams' participation in different regional and national competi-

Figure 3: Robotics Teams Increase Likelihood of Majoring in STEM More for Men



Notes: This figure reports coefficients estimated with the method from Borusyak et al. [2024] using equation 2 separately for men (left) and women (right), our event study regression, for the outcomes majoring in STEM (panel A), engineering/computer science (panel B), or computer science (panel C). All regressions include fixed effects for high school, state, and student's graduation cohort and controls for student gender, race, parental income, parental education, and student citizenship. 95% confidence intervals included; Standard errors clustered at the high school level.

tions and define a proxy for how long it takes for a team to become established: qualification for the annual national championships.[11] On average, this takes 5.4 years, which is a strikingly similar lead time observed in the dynamic specification. This is consistent with the hypothesis that robotics teams do not begin to impact student's college outcomes until they have achieved a certain degree of success.

4.2 Impacts on Pre-Major College Outcomes

In addition to the effects seen in Section 4.1, one might expect that access to FIRST robotics in high school might have intermediate effects that show up prior to deciding on a major. For some students, these extracurriculars may induce them to try computer science or engineering courses but ultimately settle on a different major. If FIRST induces students with stronger skills in these fields to try these courses, we may see that students perform better in engineering, computer sciences, and other STEM courses. Additionally, participating in FIRST may give students skills that make them more successful in their college courses. Conversely, if the students induced to try these courses due to their high school extracurriculars are less prepared, we may see worse performance among treated students. In CBII, we can observe the courses students take and the grades they received in these courses, allowing us to test how attending a high school with FIRST robotics affects course take-up and performance.

We first look at whether attending a high school with a robotics team increases the proportion of credits a student takes within a given field. Courses are categorized as falling within CIP codes, and we calculate the percent of all credits that are within STEM as a whole and in engineering and computer science individually. We then regress this on treatment using the same specification as previously. Table 4 reports the results of these regressions, adjusted as before using the Borusyak et al. [2024] method. There are limited impacts on the proportion of courses taken in these fields with one exception: men double the proportion of courses in engineering.[12] The lack of effect on STEM courses for men suggests that this is a reshuffling from other STEM courses rather than a shift from non-STEM courses.

[11] Typically, around 600-800 teams nationally qualify for championships out of the approximately 3,500 total teams, and qualification is dependent on participating in a minimum number of district events and ranking within the district.

[12] While the null effect on computer science may at first seem at odds with the impacts on majors, many of the courses taken by computer science majors are categorized as engineering classes in the database. While the average proportion of credits that are 'computer science' taken by CS majors is 2.5%, the average proportion of credits that are 'engineering' is 19.2%.

Table 4: Impacts of Robotics Team Entry on Course Choices

	(1) STEM courses	(2) CS courses	(3) Engineering Courses
Full Sample	0.001	0.0002	0.003
	(0.003)	(0.0005)	(0.003)
N	50,652	50,652	50,652
Baseline Mean:	0.234	0.006	0.047
Men	0.003	-0.0006	0.009**
	(0.004)	(0.001)	(0.004)
N	24,861	24,861	24,861
Baseline Mean:	0.250	0.073	0.007
Women	-0.001	0.001	-0.003
	(0.004)	(0.001)	(0.002)
N	25,540	25,540	25,540
Baseline Mean:	0.219	0.005	0.023
Parent BA	-0.00352	-0.0001	0.00161
	(0.00287)	(0.0006)	(0.00228)
N	67,902	67,902	67,902
Baseline Mean:	0.210	0.0132	0.0364
Parent Non-BA	-0.00160	0.0005	0.0009
	(0.00259)	(0.000728)	(0.00191)
N	78,944	78,944	78,944
Baseline Mean:	0.200	0.0109	0.0230
Family Income > 100K	-0.00404	-0.00175**	0.000815
	(0.00428)	(0.000853)	(0.00340)
N	32,648	32,648	32,648
Baseline Mean:	0.211	0.0143	0.0379
Family Income < 100K	-0.000418	-0.000922	0.00445**
	(0.00348)	(0.000898)	(0.00206)
N	42,765	42,765	42,765
Baseline Mean:	0.198	0.0128	0.0297

Notes. This table reports regression coefficients from regressions of proportion of courses in a given field on the treatment (school having a robotics team) in a bias adjusted regression using the method from Boruysak et al. implemented using Stata program did-imputation. This includes controls for high school and graduation cohort FE and clusters standard errors at the school level;* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Impacts of Robotics Team Entry on Grades

	(1) STEM GPA	(2) CS GPA	(3) Engineering GPA
Full Sample	-0.0136	-0.0767**	0.0505
	(0.0144)	(0.0337)	(0.0463)
N	92,449	17,610	24,505
Baseline Mean:	3.110	3.356	3.193
Men	-0.0272**	-0.0776*	-0.0116
	(0.0134)	(0.0420)	(0.0360)
N	42,642	9,842	16,568
Baseline Mean:	3.107	3.330	3.201
Women	-0.00824	-0.103**	0.114*
	(0.0206)	(0.0430)	(0.0654)
N	49,156	7,553	7,596
Baseline Mean:	3.113	3.390	3.177
Parent BA	-0.0363***	-0.0730**	0.0163
	(0.0134)	(0.0370)	(0.0412)
N	62,362	12,024	16,848
Baseline Mean:	3.160	3.426	3.248
Parent Non-BA	0.0347	-0.124***	-0.00377
	(0.0221)	(0.0374)	(0.0537)
N	23,115	4,476	5,855
Baseline Mean:	3.060	3.272	3.114
Family Income > 100 K	-0.0445**	-0.142***	-0.0178
	(0.0193)	(0.0498)	(0.0578)
N	31,799	6,435	8,897
Baseline Mean:	3.195	3.454	3.271
Family Income < 100 K	-0.0206	-0.0724**	0.0518
	(0.0161)	(0.0364)	(0.0362)
N	39,990	7,672	10,226
Baseline Mean	3.010	3.258	3.104

Notes. This table reports regression coefficients from regressions of average GPA within courses taken in a major type (STEM: CIP 1, 11, 14,15, 26, 27, 40, 41,51; CS or Engineering:11, 14, 15; CS: 11) on the treatment (school having a robotics team) in a bias adjusted regression using the method from Boruysak et al. implemented using Stata program did-imputation. Baseline mean is the average value of the outcome for untreated students. This includes controls for demographics, high school and graduation cohort FE and clusters standard errors at the school level;* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We next look at performance in STEM, engineering, and computer science courses. Our outcome of interest is GPA within these fields, which is calculated as the sum of all course grades weighted by credits taken within a field and then divided by the total credits taken within that major grouping.[13] Table 5 reports the results of these regressions.

Across all groups, we see that attending a high school with a robotics team is associated with students performing slightly worse in computer science classes. For the full sample, these students have a -0.08 lower GPA in these courses, which is about one-quarter of a letter grade or one-tenth of a standard deviation decline in grades. This negative association holds across all demographic groups, with slightly larger coefficients for women, students whose parents do not have a bachelor's degree, and students whose parents earn more than $100,000. For all STEM courses, we see small but statistically significant declines grades for men, students whose parents have a BA, and students whose parents earn more than $100,000. There are limited effects on performance in engineering courses, with only women seeing a marginally significant positive effect of attending a high school with a robotics team on GPAs in engineering courses.

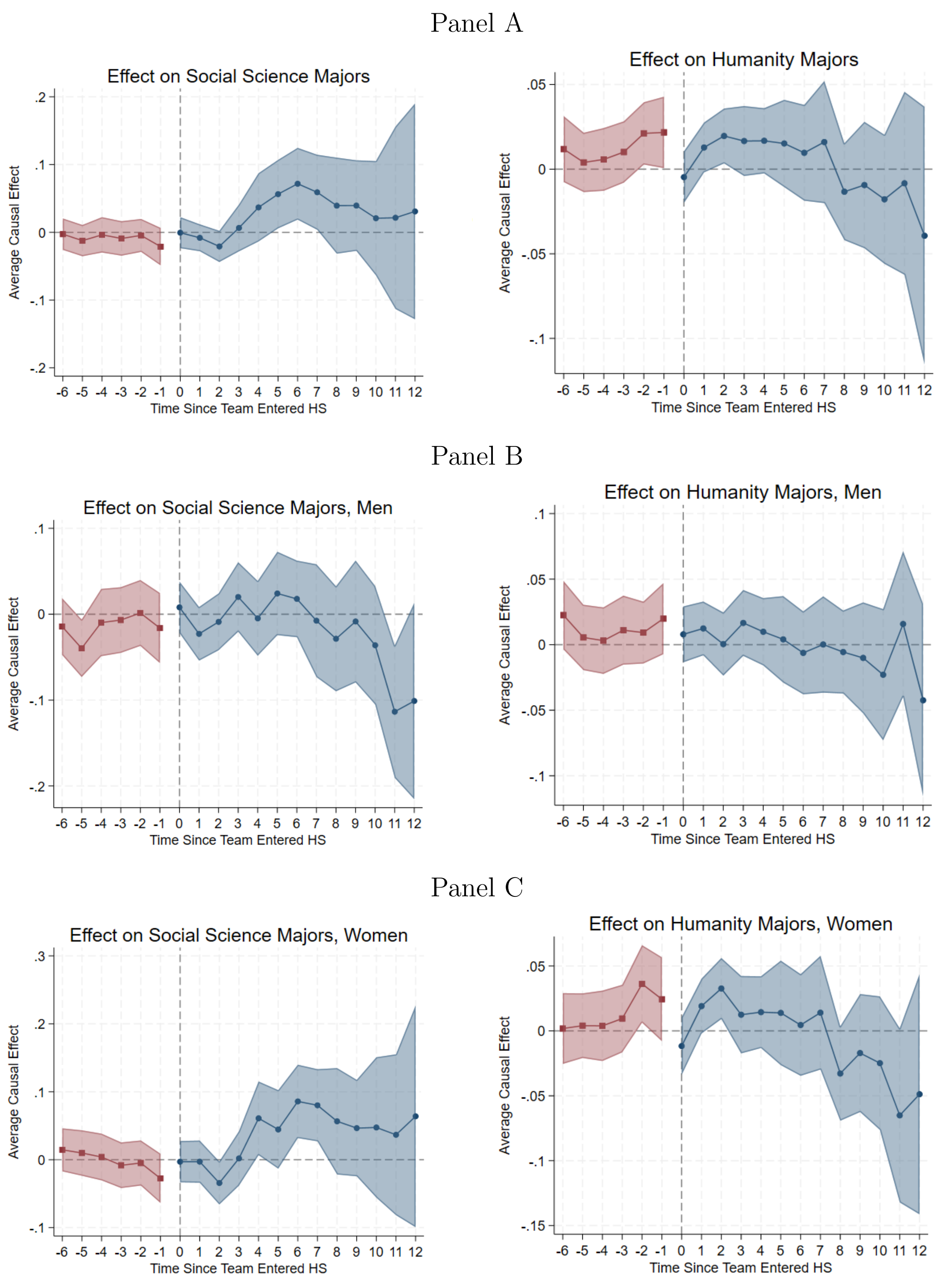
The effects on grades should be interpreted as two combined effects. First, participating in robotics in high school may encourage different students to try computer science or engineering, changing the composition of students in these courses. Second, these teams may directly effect the abilities of the students to succeed in their courses, net of compositional change. We cannot separate out these two channels, but our results are consistent with robotics clubs encouraging students with lower academic abilities to pursue computer science and STEM courses. The latter channel – direct effects on ability – are less likely for a negative effect, given that these programs increase the amount of programming experience students enter college with.

4.3 Where are Students Drawn From?

We next test whether the students entering engineering and computer science majors are drawn from alternative major categories. Figure 4 reports the effects of team entry into a high school for all students (Panel A), men (Panel B), and women (Panel C) for the outcomes of majoring in the social sciences (left) and humanities (right).

[13] GPAs are standardized across systems with grades incrementing by 0.3 grade points. That is, a 4.0 indicates an A in the course, 3.7 is an A-, 3.3 is a B+, 3.0 is a B, etc.

Figure 4: Robotics Teams Increase Likelihood of Majoring in SS and Humanities More for Women



Notes: This figure reports coefficients estimated with the method from Borusyak et al. [2024] using equation 2 separately for social science majors (left) and humanities (right), our event study regression, for the the full sample (panel A), men (panel B), or women (panel C). All regressions include fixed effects for high school, state, and student's graduation cohort and controls for student gender, race, parental income, parental education, and student citizenship. 95% confidence intervals included; Standard errors clustered at the high school level.

Team entry is associated with a 5-8 pp increase in social science majors for women, delayed by a similar number of periods as seen for the negative effects on STEM majors in figure 3. In contrast, team entry is associated with negative effects on social science majors for men, primarily concentrated many years post- team entry, corresponding to the years with the largest increases in engineering/CS majors. The effects on humanities majors are smaller and non-significant for both genders. Together, these suggest that major shifts induced by high school robotics teams are primarily shifts between STEM and social sciences, rather than humanities.

For women, the positive impact of treatment on take-up of social science majors suggests that exposure to robotics teams in high school not only reduces take-up of engineering and computer science degrees, but it directly increases the likelihood they pursue non-STEM-related majors instead. This may be driven by female students in these clubs taking on more non-robot-related parts of participating in robotics, such as the fundraising/business team which on average has a greater proportion of female students than the teams focused on building the drive train or coding [Meschede et al., 2024a]. Alternatively, these results could be driven by high school teams being less attractive or welcoming to women as extracurriculars, cementing early on a stereotype that these fields are "for" men. We explore drivers of these sex-related differences more in Section 5.

4.4 Results from the ACS

We next turn to our results from the restricted-use American Community Survey. We view this analysis as complementary with our results from the CBII data for several reasons. While we cannot observe high school of attendance in the ACS, this data allows us to observe outcomes among a nationally representative sample, unlike CBII, and also enables us to investigate whether the major substitution patterns we observe in CBII hold in an entirely different dataset and carry into the labor force. This analysis thus serves as a valuable extension and robustness check to our main findings.

We assign treatment based on year-2000 county and run regressions that are otherwise equivalent to our baseline dynamic specification (equation 2). To reduce disclosure volume, we group event-time coefficients into six bins (-6 to -4, -3 to -1, 0 to 2, 3 to 5, 6 to 8, 9 to 11). We focus on obtaining an engineering/CS major, working in a STEM-related occupation, and CPI-adjusted hourly wages as our outcome variables of interest and estimate regressions separately for men and women.

The results of this exercise are displayed in Figure 5. Several features of the results warrant discussion. First, like in our CBII data, we observe that exposure to FIRST robotics teams is associated with increased attainment in engineering/CS majors (Figures 5a and 5b). As before, we observe no evidence of differential pre-treatment trends, and the effects only phase in fully after a substantial lead time. In the first three years after team entry, men are 0.3 pp more likely to major in engineering/CS, and once the team has been established for 9-11 years, living in a county with at least one FRT increases men's likelihood of majoring in engineering/CS by 1.9 pp. While our estimates of engineering/CS major takeup among women are more precisely estimated than they were in the CBII data, they are comparable in levels and, as before, are substantially smaller than the estimated effects for men. We can reject that the effects are equal across genders in a t-test of equality of coefficients for all post-periods except 3-5 years post-team entry.

Figures 5c and 5d indicate that these effects on increased STEM major takeup carry into the labor force. For men, living in a county with a well-established FIRST robotics team is associated with a roughly 2 pp increase in the likelihood of working in STEM-related occupations post-age 25. While we also obtain positive coefficients for women, they are generally smaller in magnitude, and the presence of positive pre-treatment trends for women means that we should interpret the post-period effects with caution.[14]

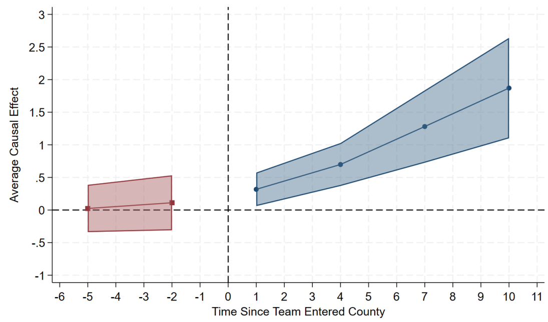
Given the increased sorting into STEM occupations we observe, a natural question is whether FIRST robotics teams go on to raise the earnings of participants. While working in STEM occupations is associated with substantially higher earnings in the labor market in the cross-section, any potential earnings effects may be substantially mitigated if these occupation substitution patterns are primarily occurring among individuals who would have otherwise gone on to work in high-earning occupations in non-STEM fields. Figures 5e and 5f suggest that this may be the case: exposure to well-established robotics teams is associated with slight *reductions* in hourly wages, but large confidence intervals and negative pre-treatment estimates render these findings, in our view, inconclusive in either direction.

We additionally investigate heterogeneity over parent income among individuals in the ACS who are linked to the 2000 Long-Form. These results are displayed in Figure B.I. As before, we observe stronger major substitution patterns among individuals with parent

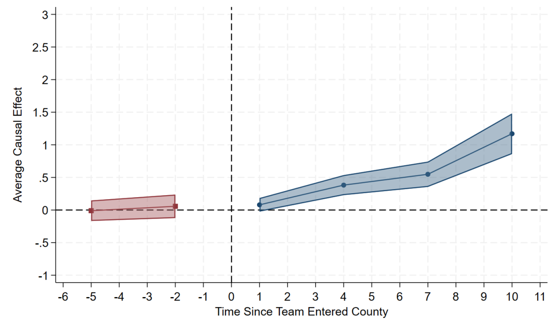
[14] While the 'trend' in the event-study may look flat albeit statistically larger than zero, the visual heuristics typically used for event studies are not appropriate evaluations of pre-trends as described in Roth [2026]. The Borusyak et al. [2024] method constructs coefficients in the pre- and post- period asymmetrically, meaning that flat pre-trends and a jump up in post-period coefficients relative to the t-1 coefficient does not indicate proof of parallel trends.

income below $100,000 (Figure B.Ia), though for both groups the results for occupational sorting and hourly wages are inconclusive. However, the fact that several key features of our baseline findings replicate in a nationally representative sample provide substantial reassurance that our results from the CBII are not being driven by a particular university or some other idiosyncratic feature of the data.

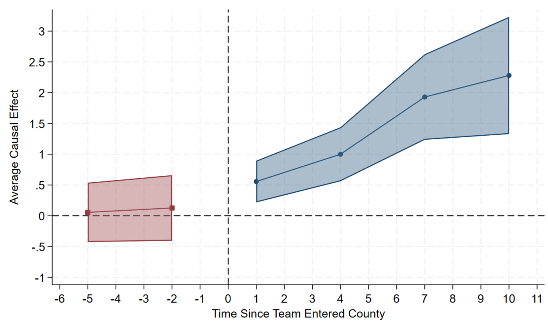
Figure 5: Impacts of Robotics Teams are Similar in the Nationally-Representative ACS Sample



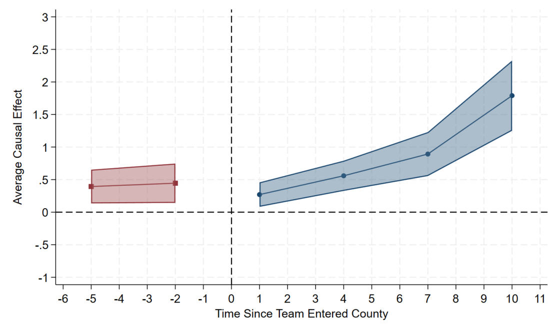
(a) Engineering/CS Major, Male



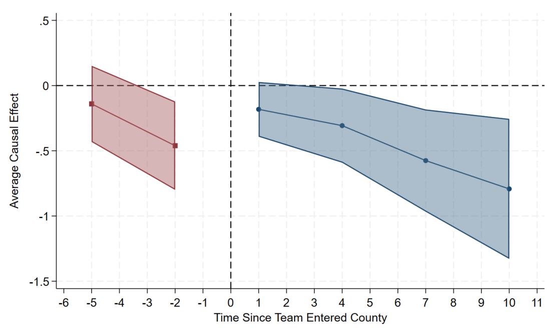
(b) Engineering/CS Major, Female



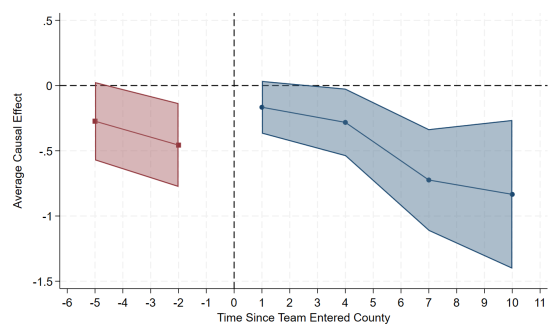
(c) STEM Occupation, Male



(d) STEM Occupation, Female



(e) Hourly Wage, Male



(f) Hourly Wage, Female

Notes: This figure reports event-study coefficients estimated with the method from Borusyak et al. [2024] using equation 2 separately for men (left) and women (right), for the outcomes majoring in engineering/computer science (panel a), working in a STEM occupation (panel b), and hourly wages (panel c). Data from 1982-1999 birth cohorts age 25 and older in 2009-2024 waves of ACS, linked to themselves via PIKs in 2000 Decennial Census. Earnings deflated to 2020 dollars using CPI. All regressions include fixed effects for county and student's graduation cohort and controls for student gender, race, and student citizenship. 95% confidence intervals included; Standard errors clustered at the county level.
CBDRB-FY26-P3098-R12992

5 Exploring Drivers of Gender Differences

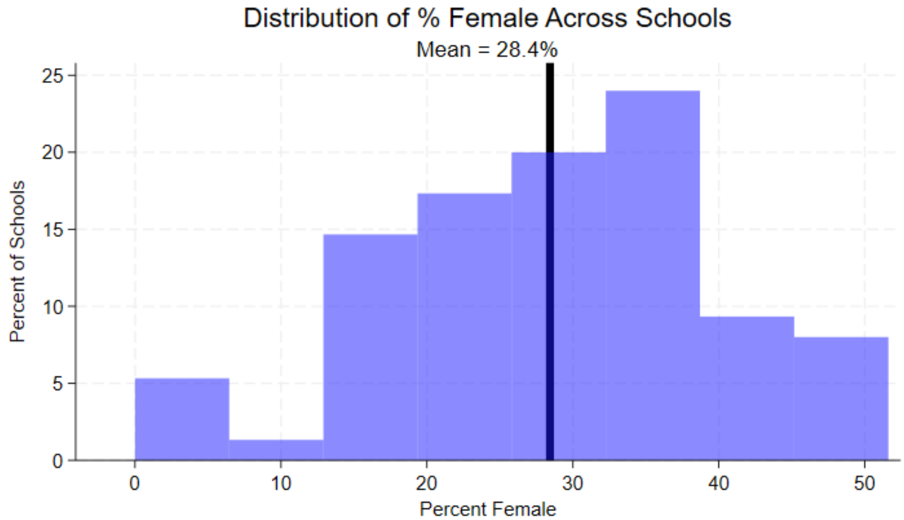
Our findings suggest that while FIRST robotics was successful in increasing students' interest in engineering and computer science degrees, it may have come at the expense of widening the gender gap in these fields. This contrasts with the findings from the FIRST Impacts longitudinal study [Meschede et al., 2024b], which followed the college trajectories of students who participated in FIRST compared to a matched sample of students who did not participate and found that FIRST participation increases STEM majors with particularly strong, positive impacts for women. How do we justify these findings?

Rather than look at the effect of *participating* in FIRST for female students who choose to be in the robotics club during high school, our findings concern the effect of being in a high school with a robotics team. In their comparisons, Meschede et al. [2024b] find that compared to women who do not choose to be on a robotics team, those who do are much more likely to major in STEM. One explanation for this is purely selection: the type of woman who chooses to be on a robotics team differs on unobservables that drive major choice, and participation in robotics clubs did not causally change their outcome. By estimating an 'intent-to-treat' effect based on entry of a robotics club into a high school rather than student's own participation, our results are not biased by selection into participation. If male students in our treated schools are more likely to participate in the robotics club than observably similar female students in our treated schools, this could be because the treatment effect works through actual participation.

Another explanation for the disparities in our findings could be that robotics teams have discouraging effects for female students who do not join the activity in high school due to the perception that robotics is a "male" extracurricular. A descriptive study of schools with robotics teams [Meschede et al., 2024a] found that robotics extracurricular participants were 68% male, whereas students in the same schools participating in other science and math extracurriculars were only 42% male. If women prefer not to be in a majority male work environment, observing the gender make-up of robotics teams in their high school may reduce their interest in engineering/computer science majors, regardless of participation.

Additionally, there may be differences in male and female students' experiences within the robotics clubs that differentially drive their likelihood of sorting into computer science or engineering majors. In interviews in Meschede et al. [2024a], women report that they felt excluded from the robot building portion of the club, saying "their male peers took over nearly all tasks relating to their team robot, including making adjustments to technical

Figure 6: Robotics Teams Are Disproportionately Male



Notes: This figure shows a histogram of the gender ratio of robotics teams based on our sample of survey respondents (N=128). The dotted vertical line represents the mean value of percent female.

details, driving the robot, and other smaller tasks." Qualitative surveys of participants at robotics competitions find that female students were less involved in programming the robot, even after controlling for prior programming experience [Witherspoon et al., 2016].

To test whether gender composition of robotics teams may drive our results, we conducted a survey of 1,000 robotics teams from our sample in which we asked respondents to report their team's gender composition by year. Teams were incentivized to participate with a chance to win one of ten $50 gift cards. 18% of invited teams started the survey and 12.4% of invited teams completed the full survey. For each participating team, we calculated the average gender ratio of students. Figure 6 shows the distribution of female participation across schools. Not a single team has more than 50% female students, and the mean proportion of women is 28%.

We next re-run our main specifications for the outcome "majoring in engineering or CS" separately for the survey sample and for teams with low female representation (the bottom tercile, or $<$ 20% female) or high female representation ($>$ 20% female).[15] If the stronger impacts in the main sample are merely driven by different participation rates, we would expect that there are positive effects for female students in high female representation

[15] We have also tested alternative cut points such as above median (28.9%). Results are of similar sign and magnitudes.

Table 6: Heterogeneity of effects, by gender composition of robotics team

	Male Students			Female Students		
	Survey Subset	<20% Female Team	>20% Female Team	Survey Subset	<20% Female Team	>20% Female Team
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.101***	0.114***	0.0971***	0.0170**	-0.0271***	0.0227**
	(0.0237)	(0.00826)	(0.0260)	(0.00847)	(0.00356)	(0.00887)
N	24,756	22,756	24,403	28,659	26,569	28,253
Baseline Mean:	0.269	0.269	0.269	0.076	0.076	0.076

Notes. This table reports regression coefficients, from regression indicators for major in computer science or engineering (CS or Engineering :11, 14, 15) on the treatment (school having a robotics team) in a bias adjusted regression using the method from Boruysak et al. implemented using did-imputation. Column 1-3 are for male students; Columne 4-6 for female students. Col. 1 and 4 are for the sub-sample of treated schools for which we have demographic information on team composition relative to never-treated schools. Col. 2 and 5 are for teams with less than 20% female participants relative to never-treated schools. Col. 3 and 6 are for teams with more than 20% female participants relative to never-treated schools. This includes controls for demographics, high school and graduation cohort FE and clusters standard errors at the school level;* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

schools and null effects for female students in low female representation schools. If there is instead a negative impact on women in low female representation schools, this is more consistent with male-dominated teams discouraging female interest.

Table 6 reports the results of these regressions. Overall, schools which responded to our survey had larger positive effects on uptake of computer science and engineering than the full sample for both genders, consistent with the fact that survey respondents are likely more established, engaged teams than non-responders. Male students who attended a high school with a robotics team that responded to the survey are 10.1 pp more likely to major in computer science or engineering than students in high schools without robotics teams, and female students are 1.7 pp more likely to choose these majors. When we split by gender composition, we see that the gender composition of the team does not significantly impact male students' likelihood of majoring in engineering or computer science but does differentially impact women. Female students on low representation teams are 2.7pp *less* likely to major in engineering or computer science, whereas women on teams with higher representation are 2.3pp *more* likely to choose these majors. This is consistent with a story in which robotics teams that are especially male-dominated have a chilling effect on women's interest in the engineering and computer science in college and helps reconcile our findings with those of Meschede et al. [2024a].

6 Conclusion

In this paper, we explore how exposure to STEM-related extracurriculars as a high school student impacts students' future degree choices in college. We leverage variation in the roll-out of FIRST robotics teams as an extracurricular at high schools across time to show, across multiple sources of data, that attending a high school that has a robotics team increases the likelihood of majoring in engineering or computer science, with larger effects for students at schools with teams that have been established for more than five years. These effects continue to matter even once individuals enter the labor force, with treated students being more likely to work in STEM-related occupations, though our ambiguous results for earnings suggest that these occupation substitution patterns may primarily be occurring among individuals who would have had high earnings in the labor market either way.

In addition to overall effects, we also see different effects by gender, with strong positive effects of take-up of engineering and computer science majors for men in high schools with robotics teams. In contrast, women's take-up of these majors increases less than similarly treated men or even decreases at times. This suggests that early exposure to robotics and technology-based extracurriculars could actually increase the gender gap in participation in engineering, perhaps due to increased salience of gender gaps in the field at earlier stages.

A limitation of our study is that we cannot directly observe whether a student in our data participated in FIRST robotics or merely attended a high school with a robotics team. This means that our estimates should be interpreted as the impacts of attending a high school with a team rather than participating in the extracurricular. These results are therefore a combined effect of actual participation and the effect of seeing others participate while not participating oneself. These potential spillover effects on non-participants may have a role in explaining our null/negative effects for women in particular. Further research into how to promote STEM-related career pathways early in the life cycle without unintentionally discouraging underrepresented groups from pursuing them would likely be worthwhile.

References

J. G. Altonji, E. Blom, and C. Meghir. Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4:185–223, 7 2012. ISSN 19411383. doi: 10.1146/annurev-economics-080511-110908.

R. J. Andrews, S. A. Imberman, M. F. Lovenheim, and K. Stange. The returns to college major choice: Average and distributional effects, career trajectories, and earnings variability. *The Review of Economics and Statistics*, 9 2024. ISSN 0034-6535. doi: 10.1162/REST_A_01503. URL https://dx.doi.org/10.1162/rest_a_01503.

S. Anwar, N. A. Bascou, M. Menekse, and A. Kardgar. A systematic review of studies on educational robotics, 2019. ISSN 21579288.

P. Arcidiacono. Ability sorting and the returns to college major. *Journal of Econometrics*, 121:343–375, 7 2004. ISSN 03044076. doi: 10.1016/j.jeconom.2003.10.010.

A. Aughinbaugh. The effects of high school math curriculum on college attendance: Evidence from the NLSY97. *Economics of Education Review*, 31:861–870, 12 2012. ISSN 0272-7757. doi: 10.1016/J.ECONEDUREV.2012.06.004.

E. Blom, B. C. Cadena, and B. J. Keys. Investment over the business cycle: Insights from college major choice. *Journal of Labor Economics*, 39:1043–1082, 10 2021. ISSN 0734306X. doi: 10.1086/712611.

P. Bordon and C. Fu. College-major choice to college-then-major choice. *The Review of Economic Studies*, 82(4):1247–1288, 2015.

K. Borusyak, X. Jaravel, and J. Spiess. Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, page rdae007, 2024.

M. C. Bottia, E. Stearns, R. A. Mickelson, S. Moller, and L. Valentino. Growing the roots of stem majors: Female math and science high school faculty and the participation of students in stem. *Economics of Education Review*, 45:14–27, 4 2015. ISSN 0272-7757. doi: 10.1016/J.ECONEDUREV.2015.01.002.

T. Breda, J. Grenet, M. Monnet, and C. Van Effenterre. How effective are female role models in steering girls towards STEM? Evidence from french high schools. *The Economic Journal*, 133(653):1773–1809, 2023.

A. A. Brenøe and U. Zölitz. Exposure to more female peers widens the gender gap in stem participation. *Journal of Labor Economics*, 38(4):1009–1054, 2020.

J. Bruhn, J. Fabian, L. Gallagher, M. Gudgeon, A. Isen, A. R. Phipps, and E. L. Gallagher. Path dependence in the labor market: The long-run effects of early career occupational experience. 2025. URL http://www.nber.org/papers/w34463.

E. J. Brunner, S. M. Dougherty, and S. L. Ross. The effects of career and technical education: Evidence from the connecticut technical high school system. *The Review of Economics and Statistics*, 105:867–882, 7 2023. ISSN 0034-6535. doi: 10.1162/REST_A_01098. URL https://dx.doi.org/10.1162/rest_a_01098.

D. Card and A. A. Payne. High school choices and the gender gap in stem. *Economic Inquiry*, 59:9–28, 1 2021. ISSN 1465-7295. doi: 10.1111/ECIN.12934. URL https://onlinelibrary.wiley.com/doi/full/10.1111/ecin.12934https://onlinelibrary.wiley.com/doi/abs/10.1111/ecin.12934https://onlinelibrary.wiley.com/doi/10.1111/ecin.12934.

S. E. Carrell, M. E. Page, and J. E. West. Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 125:1101–1144, 8 2010. ISSN 0033-5533. doi: 10.1162/QJEC.2010.125.3.1101. URL https://dx.doi.org/10.1162/qjec.2010.125.3.1101.

S. R. Cohodes, H. Ho, and S. C. Robles. Diversifying the stem pipeline: Evidence from stem summer programs for underrepresented youth. *Journal of Human Resources*, 2026. ISSN 0022-166X. doi: 10.3368/JHR.1024-13934R2. URL https://jhr.uwpress.org/content/early/2026/04/01/jhr.1024-13934R2https://jhr.uwpress.org/content/early/2026/04/01/jhr.1024-13934R2.abstract.

D. Conger, A. I. Kennedy, M. C. Long, and R. McGhee. The effect of advanced placement science on students' skills, confidence, and stress. *Journal of Human Resources*, 56:93–124, 1 2021. ISSN 0022-166X. doi: 10.3368/JHR.56.1.0118-9298R3. URL https://jhr.uwpress.org/content/56/1/93https://jhr.uwpress.org/content/56/1/93.abstract.

J. J. Conlon. Major malfunction: A field experiment correcting undergraduates' beliefs about salaries. *Journal of Human Resources*, 56:922–939, 2021. URL https://ssrn.com/abstract=2965743.

J. J. Conlon and D. Patel. What jobs come to mind? Stereotypes about fields of study. *The Quarterly Journal of Economics*, page qjag031, 2026.

G. B. Dahl, D.-O. Rooth, and A. Stenberg. High school majors and future earnings. *American Economic Journal: Applied Economics*, 15(1):351–382, 2023.

R. Darolia, C. Koedel, J. B. Main, J. F. Ndashimye, and J. Yan. High school course access and postsecondary stem enrollment and attainment. *Educational Evaluation and Policy Analysis*, 42:22–45, 3 2020. ISSN 19351062. doi: 10.3102/0162373719876923.

S. M. Dougherty. The effect of career and technical education on human capital accumulation: Causal evidence from Massachusetts. *Education Finance and Policy*, 13:119–148, 3 2018. ISSN 1557-3060. doi: 10.1162/EDFP_A_00224. URL https://dx.doi.org/10.1162/edfp_a_00224.

H. Fricke, J. Grogger, and A. Steinmayr. Exposure to academic fields and college major choice. *Economics of Education Review*, 64:199–213, 2018.

S. Granato. Early influences and the choice of college major: Can policies reduce the gender gap in scientific curricula (STEM)? *Journal of Policy Modeling*, 45:494–521, 5 2023. ISSN 01618938. doi: 10.1016/j.jpolmod.2023.04.006.

A. L. Griffith. Faculty gender in the college classroom: Does it matter for achievement and major choice? *Southern Economic Journal*, 81:211–231, 7 2014. ISSN 2325-8012. doi: 10.4284/0038-4038-2012.100. URL https://onlinelibrary.wiley.com/doi/full/10.4284/0038-4038-2012.100.

A. L. Griffith and J. B. Main. The role of the teaching assistant: Female role models in the classroom. *Economics of Education Review*, 85:102179, 12 2021. ISSN 0272-7757. doi: 10.1016/J.ECONEDUREV.2021.102179.

O. Gurantz. How college credit in high school impacts postsecondary course-taking: The role of advanced placement exams. *Education Finance and Policy*, 16:233–255, 4 2021. ISSN 1557-3060. doi: 10.1162/EDFP_A_00298. URL https://dx.doi.org/10.1162/edfp_a_00298.

M. K. Humlum, A. B. Nandrup, and N. Smith. Closing or reproducing the gender gap? Parental transmission, social norms and education choice. *Journal of Population*

Economics, 32:455–500, 4 2019. ISSN 09331433. URL https://link.springer.com/article/10.1007/s00148-018-0692-1.

A. Jackson, N. Mentzer, and R. Kramer-Bottiglio. Pilot analysis of the impacts of soft robotics design on high-school student engineering perceptions. *International Journal of Technology and Design Education*, 29:1083–1104, 11 2019. ISSN 15731804. doi: 10.1007/s10798-018-9478-8.

J. J. Kemple and C. J. Willner. Career academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood. Technical report, MDRC, 6 2008. URL www.mdrc.org.

M. C. Long, D. Conger, and P. Iatarola. Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal*, 49(2):285–322, 2012.

D. Lyles, P. Shaffery, U. Guzman, S. Spielman, and B. Weiner. CU-Boulder : First release of CEEB-NCES crosswalk (v1.0.0). https://doi.org/10.5281/zenodo.4008791, 2020.

O. Malamud. Breadth versus depth: the timing of specialization in higher education. *Labour*, 24(4):359–390, 2010.

R. Mason, G. Cooper, and T. Comber. Girls get it. *ACM Inroads*, 2:71–77, 8 2011. ISSN 21532184. doi: 10.1145/2003616.2003638. URL https://dl.acm.org/doi/10.1145/2003616.2003638.

T. Meschede, Z. Haque, M. Zhang, and M. E. Warfield. Female participants' reflection on the impacts of FIRST(for inspiration and recognition of science and technology). 2024a.

T. Meschede, M. E. Warfield, M. Hoover, Z. Haque, M. Zhang, A. Melchior, and C. Burack. The FIRST longitudinal study 10 years of follow-up data final report acknowledgements, 2024b.

A. Patnaik, M. Wiswall, and B. Zafar. College majors 1. *The Routledge handbook of the economics of education*, pages 415–457, 2021.

J. Philipp. Gendered university major choice: the role of intergenerational transmission. *Journal of Population Economics*, 36:1049–1097, 4 2023. ISSN 14321475. doi: 10.1007/s00148-022-00900-6.

J. Price. The effect of instructor race and gender on student persistence in stem fields. *Economics of Education Review*, 29:901–910, 12 2010. ISSN 0272-7757. doi: 10.1016/J.ECONEDUREV.2010.07.009.

S. F. Reardon, A. D. Ho, B. R. Shear, E. M. Fahle, and J. Kalogrides, D.and Saliba. Stanford education data archive (version 5.0). https://purl.stanford.edu/cs829jn7849, 2024.

J. Roth. Interpreting event-studies from recent difference-in-differences methods. *The Japanese Economic Review*, pages 1–14, 2026.

M. van der Vleuten, E. Jaspers, I. Maas, and T. van der Lippe. Intergenerational transmission of gender segregation: How parents' occupational field affects gender differences in field of study choices. *British Educational Research Journal*, 44:294–318, 4 2018. ISSN 1469-3518. doi: 10.1002/BERJ.3329. URL https://onlinelibrary.wiley.com/doi/full/10.1002/berj.3329https://onlinelibrary.wiley.com/doi/abs/10.1002/berj.3329https://bera-journals.onlinelibrary.wiley.com/doi/10.1002/berj.3329.

A. Welch and D. Huffman. The effect of robotics competitions on high school students' attitudes toward... *School Science and Mathematics*, 111:416–424, 2011. ISSN 0036-6803.

M. Wiswall and B. Zafar. How do college students respond to public information about earnings? *Journal of Human Capital*, 9:117–169, 2015a. URL http://graphicsweb.wsj.com/documents/NILF1111/.

M. Wiswall and B. Zafar. Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82:791–824, 8 2015b. ISSN 1467937X. doi: 10.1093/restud/rdu044.

E. B. Witherspoon, C. D. Schunn, R. M. Higashi, and E. C. Baehr. Gender, interest, and prior experience shape opportunities to learn programming in robotics competitions. *International Journal of STEM Education*, 3:1–12, 2016.

B. Zafar. College major choice and the gender gap. *Journal of Human Resources*, 48:545–595, 2013.

A Standard Two-Way Fixed Effects Specification

While our preferred specification uses the Borusyak et al. [2024] imputation method, we also test whether our results hold using the standard two-way fixed effects method. In Tables A.I, A.II, and A.III, we report present results from regressing indicators for majoring in STEM, engineering or computer science, and computer science alone on indicators for the treatment in our static difference-in-difference specification described in equation 1. We run this first for the full sample (col. 1 and 2.) and then allowing the treatment to vary by gender (col. 3 and 4), parental income (col. 5 and 6), and parental education (col. 7 and 8).

Table A.I: Impacts of Attending a HS with a Robotics Team: STEM Major

	(1) Overall	(2) Overall	(3) Gender	(4) Gender	(5) Income	(6) Income	(7) Parent BA	(8) Parent BA
Treated	0.002	0.006	0.014**	0.017**	0.003	-0.002	-0.014*	-0.009
	(0.005)	(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.008)	(0.010)
Treated X Female			-0.026***	-0.021***				
			(0.007)	(0.008)				
Income>100K X Treated					0.012	0.015*		
					(0.008)	(0.008)		
Parent BA X Treated							0.021***	0.020**
							(0.008)	(0.009)
Controls		X		X		X		X
R2	0.089	0.130	0.107	0.130	0.100	0.130	0.090	0.130
N	171747	106293	171747	106293	171747	106293	171747	106293

Table Notes: This reports regression coefficients for a regression of an indicator variable for majoring in STEM (defined as CIP codes 1, 11, 14, 15, 26, 27, 40, 41, and 51) on an indicator for being treated, along with interactions with indicators for gender, indicators for family income of over 100K income, and indicators for at least one parent having a bachelor's degree. Mean of the outcome variable is 0.325 for untreated students (N=137,294). Treated means that the student attended a school that had a robotics team in the previous 4 year before college first term. All regressions include fixed effects for high school, state, and student's graduation cohort. Odd columns include controls for student gender, race, parental income, parental education, and student citizenship. Standard errors in parentheses clustered at the high school level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In all specifications, there is no significant impact of attending a high school with a robotics team on major choice in the full sample. However, this null result hides substantial heterogeneity in impacts by demographic groups with particularly strong differences by gender. We see a significant positive effect of the treatment on male students on likelihood of choosing a technical major. Male students whose high school has a robotics team are 1.7 pp (2.3%) more likely to major in STEM and 1.5 pp (9.3%) more likely to major in engineering or computer science. In contrast, there is a negative coefficient on the interaction between being female and the treatment variable, indicating that attending a high school with a robotics team widens the gender gap in likelihood of majoring in a technical field. Women

Table A.II: Impacts of Attending a HS with a Robotics Team: Engineering or CS Major

	(1) Overall	(2) Overall	(3) Gender	(4) Gender	(5) Income	(6) Income	(7) Parent BA	(8) Parent BA
Treated	0.000	0.003	0.014**	0.016**	-0.004	-0.009	-0.013**	-0.015*
	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Female X Treated			-0.030***	-0.025***				
			(0.006)	(0.007)				
Income>100K X Treated					0.019***	0.022***		
					(0.007)	(0.007)		
Parent BA X Treated							0.018***	0.024***
							(0.007)	(0.007)
Controls		X		X		X		X
R2	0.096	0.162	0.151	0.163	0.099	0.163	0.097	0.163
N	171747	106293	171747	106293	171747	106293	171747	106293

Table Notes: This reports regression coefficients for a regression of an indicator variable for majoring in Computer Science or Engineering (defined as CIP codes 11, 14, 15) on an indicator for being treated, along with interactions with indicators for gender, indicators for family income of over 100K income, and indicators for at least one parent having a bachelor's degree. Mean of the outcome variable is 0.152 for untreated students (N=137,294). Treated means that the student attended a school that had a robotics team in the previous 4 year before college first term. All regressions include fixed effects for high school, state, and student's graduation cohort. Odd columns include controls for student gender, race, parental income, parental education, and student citizenship. Standard errors in parentheses clustered at the high school level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

from high schools with robotics are 2.1 pp less likely than men from similar high schools to major in STEM. If we calculate the treatment effect for women (i.e., the sum of the two coefficients), there is a null effect on STEM major take-up for women. For engineering and computer science, these disparities are even more pronounced with the treatment effect indicating reductions in major take-up relative to other women in high schools without robotics teams. For example, there is a null effect of the treatment on men's take-up of computer science specifications in the specification including demographic controls, but there is a significant, negative effect for women. This indicates that women are approximately 1.6 pp (36%) less likely to major in computer science if their high school has a robotics team than in the absence of a team.

We also see disparities by family background, though these differences are less stark and only statistically significant at standard levels for engineering and computer sciences majors. Treated students whose parents make more than \$100,000 per year are 2.2 pp (11.7%) more likely to major in engineering or computer science and 1.2 pp (24.3%) more likely to major in computer science alone than those whose high school doesn't have a robotics team, whereas students with lower earning parents have a null effect. This interaction with parental income indicates that these programs widen the gap in take-up of STEM majors by parental income. Similar patterns are seen if we measure student socioeconomic backgrounds using parental

Table A.III: Impacts of Attending a HS with a Robotics Team: CS Major

	(1) Overall	(2) Overall	(3) Gender	(4) Gender	(5) Income	(6) Income	(7) Parent BA	(8) Parent BA
Treated	-0.001	-0.004	0.009**	0.005	-0.008**	-0.010**	-0.007**	-0.012**
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Female X Treated			-0.020***	-0.016***				
			(0.004)	(0.004)				
Income>100K X Treated					0.011***	0.012***		
					(0.004)	(0.004)		
Parent BA X Treated							0.009***	0.010***
							(0.004)	(0.004)
Controls		X		X		X		X
R2	0.063	0.093	0.081	0.093	0.071	0.093	0.063	0.093
N	171747	106293	171747	106293	171747	106293	171747	106293

Table Notes: This reports regression coefficients for a regression of an indicator variable for majoring in Computer Science (defined as CIP codes 11) on an indicator for being treated, along with interactions with indicators for gender, indicators for family income of over 100K income, and indicators for at least one parent having a bachelor's degree. Mean of the outcome variable is 0.041 for untreated students (N=137,294). Treated means that the student attended a school that had a robotics team in the previous 4 year before college first term. All regressions include fixed effects for high school, state, and student's graduation cohort. Odd columns include controls for student gender, race, parental income, parental education, and student citizenship. Standard errors in parentheses clustered at the high school level; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

education rather than income.

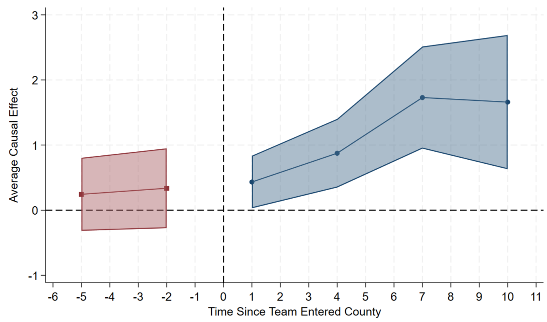
B Supplementary Tables and Figures

Table B.I: Summary Statistics of ACS Sample

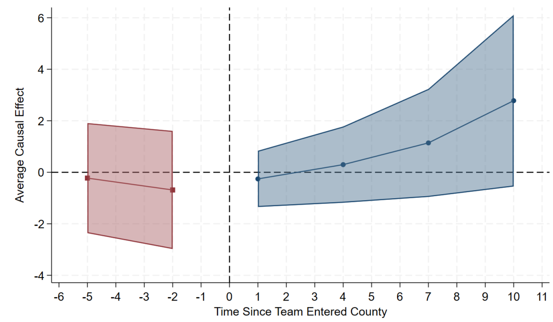
Variable	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
White	82.0	82.8	81.3	88.5	83.5
	(38.3)	(37.7)	(38.9)	(31.8)	(37.0)
Black	10.0	9.2	10.7	3.6	9.1
	(29.9)	(28.9)	(30.9)	(18.6)	(28.8)
Asian	3.7	3.8	3.6	4.5	2.8
	(18.9)	(19.1)	(18.7)	(20.7)	(16.5)
STEM BA	12.7	12.9	12.5	20.0	11.5
	(33.3)	(33.5)	(33.1)	(40.0)	(31.9)
Engineering/Compsci BA	4.5	7.5	1.5	7.4	3.9
	(20.6)	(26.3)	(12.1)	(26.1)	(19.4)
4-Yr College Degree	45.7	41.0	50.4	72.4	39.8
	(49.8)	(49.1)	(50.0)	(44.6)	(48.9)
Enrolled in Grad School	4.5	3.7	5.3	7.2	4.0
	(20.8)	(18.9)	(22.4)	(25.7)	(19.5)
Advanced Degree	13.7	10.7	16.7	25.0	10.9
	(34.4)	(30.9)	(37.3)	(43.3)	(31.2)
Employed	81.8	84.9	78.7	86.5	81.3
	(38.5)	(35.7)	(40.9)	(34.0)	(38.9)
STEM Occupation	7.9	11.7	4.1	11.6	7.2
	(26.9)	(32.1)	(19.9)	(32.1)	(25.8)
Annual Labor Income ($1,000)	43.7	51.0	36.6	60.4	39.7
	(52.1)	(59.6)	(42.3)	(70.1)	(44.4)
Hourly Wage	24.7	26.1	23.3	30.9	23.0
	(19.7)	(20.6)	(18.5)	(23.2)	(18.0)
Sample	All	Men	Women	Par Inc>100k	Par Inc<100k
N	2990000	1470000	1520000	61700	315000

Notes: Table presents summary statistics of ACS sample. Standard deviations in parentheses. Data from 1982-1999 birth cohorts age 25 and older in 2009-2024 waves of ACS, linked to themselves via PIKs in 2000 Decennial Census. Parent income split conducted for individuals linked to 2000 Long-Form Decennial. Earnings deflated to 2020 dollars using CPI. Sample sizes rounded for disclosure avoidance purposes. CBDRB-FY26-P3098-R12992.

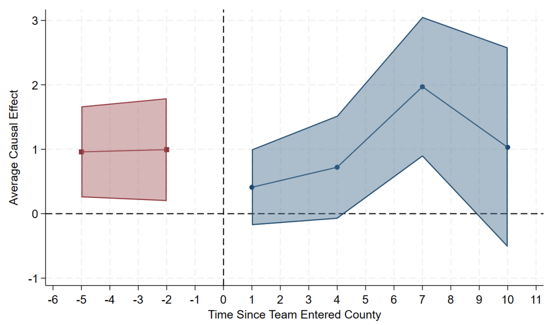
Figure B.I: ACS Results by Parent Income



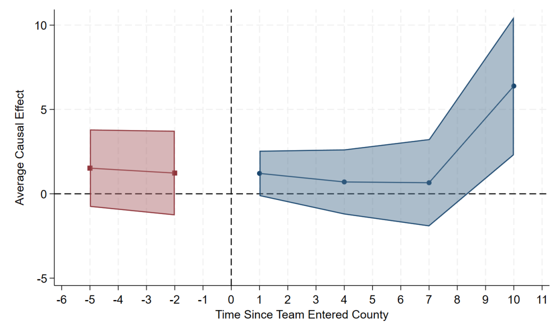
(a) Engineering/CS Major, < \$100k



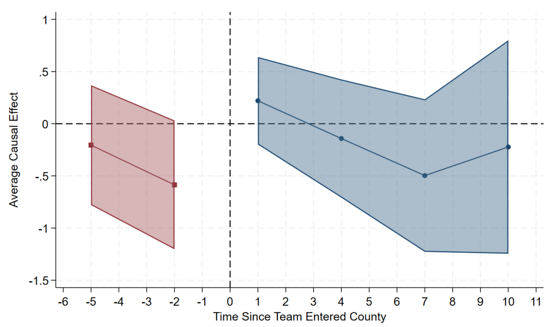
(b) Engineering/CS Major, > \$100k



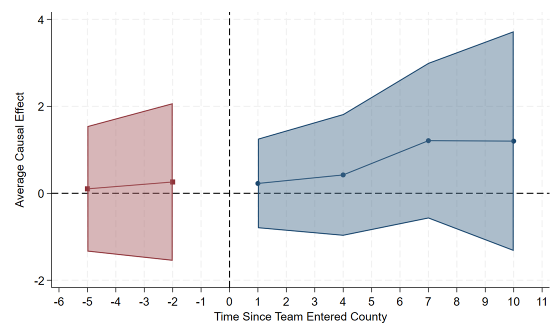
(c) STEM Occupation, < \$100k



(d) STEM Occupation, > \$100k



(e) Hourly Wage, < \$100k



(f) Hourly Wage, > \$100k

Notes: Figures present event-study estimates using Borusyak et al. [2024] method of exposure to FIRST robotics teams on outcome variables of interest. Dotted lines show 95% confidence intervals. Data from 1982-1999 birth cohorts age 25 and older in 2009-2024 waves of ACS, linked to themselves via PIKs in 2000 Decennial Census Long-Form. Parent income from 2000 Long-Form. Earnings deflated to 2020 dollars using CPI. CBDRB-FY26-P3098-R12992