Influencing Girls' Educational Choices through Social Media Interventions: Evidence from Peru*

Marcos Agurto
Universidad de Piura

Sudipta Sarangi
Virginia Tech University

Danila Serra
Texas A&M University

August 2025

PRELIMINARY DRAFT - NOT FOR CIRCULATION

Abstract

Although social media plays a central role in teenagers' lives, there is limited evidence on its potential to shape youths' preferences, aspirations and educational choices. We conduct a randomized controlled trial in Peru to evaluate the impact of a role model intervention delivered via Instagram, complemented by an information campaign targeting teachers through WhatsApp. The Instagram intervention exposed final-year high school girls to short video reels created by female engineering students, depicting their daily academic and social experiences on campus. Girls were encouraged to follow and engage with the content over six weeks. While the interventions initially reduced self-reported interest in engineering, particularly among girls who were not top performers in math, we find that, three months post-intervention, high-achieving girls in math were significantly more likely to apply for a scholarship to study engineering, especially when their teachers were also targeted. Most notably, the Instagram intervention led to a large increase in girls' overall college enrollment, regardless of major. Our analysis of underlying mechanisms suggests that exposure to the social media content enhanced girls' confidence in their ability to succeed in higher education and positively influenced their perceptions of peers' likelihood of enrolling in university. The add-on teacher intervention appeared to amplify the impact of the Instagram campaign for top-performing girls in math, while attenuating it for their lower-performing peers. Our study demonstrates the promise of leveraging widely-used social media platforms to engage young women and reshape their perceptions of higher education.

Keywords: Social media, education, gender, Peru, experiment

*We acknowledge funding form the Consorcio de Investigación Económica y Social, Peru. The study is preregistered in the AEA Registry with ID AEARCTR-0012871.

1 Introduction

Social media is a central part of teenage life across the globe. In the United States, teenagers spend over four hours a day on social media, with platforms like Instagram, Snapchat, and TikTok used daily by more than 70 percent of those aged 15 to 17 (Anderson et al., 2023). In Peru, where we conduct our study, the rapid expansion of broadband infrastructure over the last decade (More Sánchez and Argandoña Martinez, 2020) has similarly fueled widespread adoption of social media apps, including Instagram, especially among younger users.[1] While there is active public debate and a growing body of research on the possible effects of smartphone use on teenagers' mental health and academic performance, with mixed findings (Beland and Murphy, 2016; Kessel et al., 2020; Kreski et al., 2021), there is far less evidence on whether and how these platforms can influence adolescents' preferences, aspirations, and educational choices.

In this paper, we explore how social media can be leveraged to influence educational choice of teenage girls. We designed an Instagram intervention to increase interest in STEM post-secondary education among female high school students in Peru. Despite significant progress in women's access to higher education, female representation in STEM fields, particularly engineering, remains low in many countries. In Peru, only 30% of engineering applicants are women, and less than 10% of women in college major in engineering. While in-person exposure to female teachers and role models (Lim and Meer, 2020; Sevilla et al., 2023; Agurto et al., 2023; Breda et al., 2023; Porter and Serra, 2020) has shown promise, such interventions digital platforms, and especially social media, offer a compelling alternative for reaching adolescents at scale and at low cost.

We employ a randomized controlled trial to examine how "Instagram role models" targeting high school girls can impact post-secondary educational preferences and college enrollment choices. Specifically, we created an Instagram account called "My Days in University" (Mis Dias En La U) and invited final year female students in randomly selected high schools to follow the account. These students were encouraged to engage with the content by liking a series of short videos posted on the account by female engineering college students over a total of 6 weeks. The short videos/reels documented the female students' experiences in college, while studying engineering, e.g., going to class, studying for homework, completing projects and interacting socially with other students.

Students' educational choices are also shaped by the attitudes and support of influential adults in their environment. The encouragement of teachers is particularly important when girls' preferences challenge traditional gender roles or diverge from family expectations. Prior research has shown that gender stereotypes held by teachers can dampen girls' academic confidence and contribute to their under-representation in STEM majors (Carlana, 2019). To address this, we complemented the Instagram intervention with a second component targeting teachers. We implemented an information campaign via WhatsApp, the most widely used messaging app in Peru, designed to raise teachers' awareness of gender bias and encourage them to

[1] A recent survey suggests that two-thirds of adults in Peru use Instagram; however, accurate statistics for teenagers are not available. See https://www.statista.com/forecasts/1409959/social-network-usage-by-brand-in-peru.

actively support female students interested in STEM.

The study involves 73 high schools in Peru, randomly divided into two treatment groups (48 schools combined) and one control group (25 schools). In the treatment schools, we implemented the role model intervention on Instagram. A total of 9 short videos ("reels") were shared by female engineering students over 6 weeks; after that, no additional content was posted but the account remained online, with the the videos still visible to account followers. In about half of the treatment schools (23 schools), we added the WhatsApp intervention: 3 short videos were sent via WhatsApp to the head teachers of the graduating student cohorts. These videos featured engineering professors discussing the gender imbalance within the engineering field and emphasizing the desirability of increased female representation in this domain. The professors encouraged teachers to actively contribute to reducing the gender imbalance by challenging and dismantling the gender stereotypes associated with the engineering major. This dual approach aimed to leverage the reach of social media while also targeting key "influencers," the head teachers, to reinforce the impact of the role model intervention.

While we initially aimed to target all final-year high school students in the sampled schools, in collaboration with school principals and head teachers, the requirement to obtain parental consent for social media content exposure and collect national ID numbers for future tracking led to a final sample to 547 girls, representing approximately 20 percent of the graduating class. Although the intervention and primary analysis focus on girls, we also surveyed a small sample of boys to enable gender gap comparisons in educational choices, bringing the total sample to 787 final-year students. All participants filled in a baseline online survey (in July and August 2023), where we also recorded the unique student identification number, which allows following up the students if/when they enroll in college.

We implemented a follow-up survey with the study participants immediately after the implementation of the 6 weeks of Instagram-based role model program, between October and November 2023. This survey measured students' preferences regarding different educational and career paths, as well as their perceptions of their parents', head teacher's and best friends' preferences. About two months later, in early January 2024, after most students intending to attend college had completed their applications,[2] we offered all students the opportunity to apply for a small scholarship, which would be randomly awarded to one applicant, contingent on enrollment in an engineering major the following year. This scholarship application process served as a revealed preference measure for interest in engineering, as applying required filling out paperwork, making it more likely that only students seriously considering an engineering major would apply. Given the small size of the scholarship (about $750),[3] we do not expect it to have significantly shifted students' preferences. Finally, in December 2024, we collected administrative data on actual students' college enrollment, about one year after students' high school graduation.

Our findings indicate that the interventions had an immediate negative impact on girls' self-reported

[2] College applications in Peru are typically submitted between October and December of the year preceding college enrollment. Universities have two terms, March to June, and August to November. Most enrollment by freshmen happens in the first term.

[3] We anticipated the 2775 soles scholarship to cover for about 4 months of room and board in a public university, or for enrollment fees and 1 month of tuition in a private elite university.

preferences for engineering, and no impact on the decision to apply for the scholarship. However, these aggregate findings mask important heterogeneities driven by girls' math ability. Specifically, the negative treatment effect on self-reported educational preferences is driven by girls who are not top performers in math. For these students, the interventions appear to have reinforced their intention to *avoid* studying engineering. Additionally, we observe no impact on their decision to apply for the scholarship, with only a small percentage choosing to do so. In contrast, a large and significant positive impact on scholarship applications is observed among top-performing girls in math, particularly when the role model intervention is combined with the teacher information campaign.

The administrative data on college enrollment show evidence of a large and positive impact of the interventions on college enrollment for all girls, regardless of major. Specifically, the Instagram intervention increases girls' likelihood of being enrolled in college one year after graduating from high school by 20 percentage points, i.e., a 75 percent increase from the 26.5 percent enrollment rate observed in the control group. We do not find evidence of an impact, however, of the interventions on enrollment in engineering, but we may be underpowered to detect such an effect. Our data also provide suggestive evidence that the teacher information intervention reinforced the impact of the role model Instagram intervention for girls who were top performing in math, while reducing its effect for not-top-performing girls. Our analysis of mechanisms suggests that the Instagram intervention operated primarily by increasing girls' confidence in their ability to complete college (but not to pursue engineering studies), and by shifting their beliefs about the likelihood that their best friends would attend college.

We draw two key takeaways from these results. First, the Instagram-based role model videos may have reshaped girls' perceptions (or misperception) of university life and their potential for success in higher education, likely leading to the overall increase in college enrollment. These effects may have been further amplified by changes in peers' educational choices through social influence within friendship networks, and by the Instagram algorithm itself, which could have exposed treated girls to similar content and accounts created by other college-aged women, reinforcing the messages delivered through our intervention. Second, information interventions targeting teachers and aimed at dismantling stereotypes regarding women in STEM may facilitate or hinder this process. For high-achieving girls, the teacher intervention seemed to reinforce the positive effects of the role model intervention, boosting their likelihood of enrolling in college. However, for girls who are not top performers in math, the teacher intervention may have unintentionally attenuated the impact of the Instagram intervention on the intention to pursue higher education, possibly due to perceived academic limitations or lower expectations.

This study contributes to several bodies of work. First, we add to research on the factors that influence youths' educational preferences and choices, particularly the role of information, advising, and role models in shaping girls' decisions to pursue education and careers that are prone to gender stereotyping. Prior studies have shown that exposure to relatable role models can influence girls' aspirations, gender attitudes, self-confidence and perceptions of different fields of study and careers, often leading to changes in educational outcomes (Agurto et al., 2023; Breda et al., 2023; Di, 2024; Kipchumba et al., 2024; Porter and Serra, 2020).[4] Our study extends this literature by showing that scalable role model interventions delivered via a widely used social media platform can shift college enrollment decisions among high school girls. However, shifting preferences toward engineering appears more challenging and may require the buy-in of parents, particularly in low-income contexts where parental concerns about gender norms and women's safety may weigh heavily on students' educational choices.[5]

We also contribute to the growing literature on the role of teachers in reinforcing or dismantling gender stereotypes, which can shape girls' academic performance and their decision to pursue male-dominated fields of study. For example, Alan et al. (2018) find that girls exposed to teachers with more traditional gender attitudes perform worse academically, while Carlana (2019) shows that teachers' gender stereotypes contribute to the gender gap in math performance and track choice. Moreover, Lavy and Megalokonomou (2024) provide evidence of the long-term impacts of high school teachers' gender bias on girls' performance in university admission exams and their choice of university field of study.[6] Our findings add to this body of work by suggesting that interventions explicitly targeting high-achieving girls, for instance to encourage their choice of a male-dominated field of study like engineering, may induce teachers to reallocate their time and attention toward these students. While such reallocation could strengthen the impact of the intervention for high-achieving girls, it may also inadvertently reduce support for girls who are not top performers, possibly leading to discouragement and lower educational attainment.

Finally, we contribute to the literature in the economics of social media. Empirical studies in this space have primarily examined the impact of social media on political polarization, access to information, and misinformation (Aridor et al., 2024a,b). A smaller body of work has investigated the welfare effects of social media (Allcott et al., 2020; Bursztyn et al., 2023; Mosquera et al., 2020). These studies typically recruit participants through social media platforms, expose them to different content or incentives, and measure outcomes via follow-up surveys administered on the same platform or by tracking platform usage patterns. Our contribution to this literature is threefold. First, we recruit participants outside of the platform and measure real-life outcomes, moving beyond self-reported or platform-based metrics; in particular, we focus on impacts on post-secondary educational choices. Second, we specifically target teenagers and use Instagram, a platform widely used by younger generations but not targeted by previous social media interventions.[7] Third, we introduce a novel approach to studying the effects of social media exposure on teenagers' preferences and behaviors: we created a dedicated Instagram account, invited participants to consume its content, and measured its effects on behaviors outside of the social media environment.

Overall, our findings highlight the potential of leveraging social media platforms that are heavily used

[4] For a review of studies testing the impact of role models on a variety of outcomes in low-income countries, see Serra (2025).

[5] Other studies have tested more time-intensive mentoring or advising interventions targeting high school students. See, for instance, Barr and Castleman (2025), Bettinger and Evans (2019), Carrell and Sacerdote (2017).

[6] Related studies have provided evidence of the impact of female teachers on women's choice of male-dominated field of study. See, for instance, Bettinger and Long (2005), Lim and Meer (2020) and Carrell et al. (2010).

[7] Most social media experiments to date have been conducted on Facebook and Twitter, as documented in Aridor et al. (2024a), but these platforms are far less popular among teenagers, who predominantly use Instagram, TikTok, and YouTube.

by teenagers to influence their educational choices. As internet access and smartphone use continue to expand in both high-income and low-income settings, social media offers a low-cost and scalable channel to reach young people at a formative stage of their decision-making, possibly more directly and frequently than traditional school-based or community outreach programs. Changing girls' preferences for engineering, however, may require more targeted interventions and the active involvement of parents. Our results also highlight the important role of teachers: while they can reinforce the messages delivered through social media, interventions must be designed with care to avoid inadvertently diverting time and attention away from students who are not the primary focus of the program.

2 Context

The education system in Peru comprises six years of primary schooling followed by five years of secondary schooling, totaling 11 years. Attendance is compulsory from ages 5 to 16. At the secondary level, nearly 2.5 million students are enrolled in approximately 15,000 high schools across the country's 25 administrative regions. Secondary education is subject-based, with different teachers assigned to each subject, and the academic year runs from March to December.[8] About 76% of secondary students attend public schools managed by the central government, while the remainder are enrolled in for-profit and non-profit private institutions. All schools follow the national curriculum established by the Ministry of Education, which makes no distinction between students intending to pursue STEM or non-STEM careers at any stage of basic education.

In both public and private schools, each grade is generally divided into several sections. Students within the same section form close-knit social networks, usually advancing together from the first to the final year of high school and taking all their classes as a group. This structure fosters friendships that are predominantly formed within the section. Each section has a head teacher, or "tutor," who interacts closely with all students and provides advice on academic matters as well as future educational and career choices.

Our study focuses on high schools in six northern regions of Peru: Piura, La Libertad, Cajamarca, Ancash, Lambayeque, and Tumbes. See Figure A1 in Appendix for a visualization of these regions in Northern Peru. The Piura region, home to the main campus of the Universidad de Piura (UDEP), one of the major universities in the North of Peru, accounts for about 65 percent of the sample schools. Secondary school enrollment in our study area is approximately 96.3%, close to the national average of 97%. Moreover, 50.2% of high school students in our sample are in the grade level corresponding to their age, slightly below the national average of 52.8%. In terms of post-secondary education, college enrollment among individuals aged 17–24 in the study regions is around 31.5%, below both the national average of 33% and the coastal-region average of about 35%.[9]

[8]The national curriculum includes Mathematics, Communication, Foreign Language, Art, History, Geography, Economics, Civics, Social Skills, Physical Education, Religious Education, Science, Technology, and Environmental Studies.

[9]Fuente: INEI – *Indicadores de educación según departamentos*, 2013–2023. https://cdn.www.gob.pe/uploads/document/file/7043096/6061323-peru-indicadores-de-educacion-segun-departamentos-2013-2023.pdf?v=1728072117

University admissions in Peru are decentralized, with each institution designing and administering its own process. Public universities typically rely on competitive entrance examinations specific to each major, while some private institutions offer additional admission channels.[10] The admissions process usually takes place between September and March, and the academic year typically begins in late March.

While STEM careers in Peru encompass various disciplines, engineering overwhelmingly dominates students' STEM choices. In 2016–2017, roughly 93% of the 417,000 applicants to STEM programs applied to engineering majors. As in many countries in Latin America and elsewhere, female participation in STEM, and in engineering in particular, remains low. During that period, only 30% of STEM applicants were women,[11] and just 19% of female applicants nationwide chose engineering as their intended major, compared to 46% of male applicants. The gender gap is even more apparent when looking at enrollment and graduation rates. Recent OECD data reveal that only 9% of women as opposed to 29% of men entered college to pursue a Bachelor degree engineering in 2023, the latest year with data availability and the year of our intervention.[12]

3 The Social Media Interventions

We recruited 73 high schools from 6 regions of Northern Peru to participate in the study. This was made possible by the ongoing relationship between high schools and the Universidad de Piura (the home base of one of the authors). We randomly assigned the schools into three groups: a Control Group (C), an Instagram Intervention (T1) group and an Instagram & WhatApp Intervention (T2) group. The school-level randomization was stratified by municipality, by whether the school was co-ed or girls-only (about 86 percent of schools in the sample are co-ed), and by the Socio-Economic status (*high or low*) assigned to the school by the Universidad de Piura. We describe interventions in detail below.

3.1 Design and Implementation

The Instagram intervention involved the creation of an account designed to showcase the daily lives of women studying engineering in college. Girls in the treatment group were invited to follow the account and engage with the content, which was posted weekly over a six-week period. The account was managed by female engineering students at the Universidad de Piura, who designed the logo (see Figure 1) and produced short video reels – typically 5 to 10 seconds long – depicting their experiences on campus. The reels followed the typical Instagram format, incorporating music and humorous or eye-catching text. Importantly, the college women were given full creative freedom to decide what content to create and share in order to inspire high

[10] For example, certain private universities grant direct admission to students ranked in the top third of their graduating class.

[11] STEM fields include Biology, Mathematics, Statistics, Engineering, Physics, and Chemistry. Medicine and related health sciences are excluded from this classification; in medical programs, women are the majority, representing about 70% of enrolled students.

[12] The OECD data can be downloaed here: https://gpseducation.oecd.org/CountryProfile?primaryCountry=PER&treshold=5&topic=EO

school students to follow in their footsteps.

Some reels featured photos of and references to female scientists who had influenced the students' academic journeys. Others showed scenes of campus life, such as walking through university buildings, studying alongside other women, taking exams, and discussing coursework with professors. The overall goal was to present relatable peer role models from similar backgrounds, inspire high school viewers, and make college life, particularly in engineering, appear more accessible, enjoyable, and compatible with their identities. By doing so, the intervention aimed to dismantle stereotypes about women in higher education, and specifically about their under-representation in engineering.

In order to avoid contamination of the control group, the account was set as private and only girls assigned to the treatment group were provided access, if they requested. This was made possible by the implementation of the baseline survey before the intervention. Nearly 90 percent of the surveyed girls had an Instagram account at baseline.[13] This allowed the account managers to verify the treatment-group nature of the girls requesting access to the account. Any possible requests coming from boys were also rejected.

The Instagram account was promoted by the head teachers of final years students in the treatment group during school hours. Its handle was posted in class on the board, and head teachers encouraged girls to follow the account. While we aimed to target all girls in the treatment schools with the Instagram intervention, only the girls for whom we obtained parental consent during the baseline survey were granted access to the Instagram account. In line with typical engagement strategies on social media, followers received small incentives for liking the weekly content, such as randomly distributed school supplies, including pens, markers, and small bags.

The only difference between the *Instagram (T1)*and the *Instagram & WhatsApp (T2)* treatment groups is that the schools randomized into the latter also received an intervention targeting the head teachers of the graduating class. Specifically, we sent these teachers three short videos (each approximately 30 seconds long) via WhatsApp. The videos featured three engineering professors – two women and one man – speaking directly to teachers. In their messages, the professors emphasized the under-representation of women in engineering, affirmed the capabilities of women to succeed in the field, and encouraged teachers to actively support and motivate more girls to pursue engineering studies.[14]

3.2 Timeline, Data Collection and Outcomes of Interest

In July 2023, we implemented a baseline online survey of final-year students in all sampled high schools. The survey was promoted in class by teachers, who provided students with the survey link. It collected information on students' educational aspirations and expectations, beliefs about their parents' educational preferences, perceived support from teachers, prior-year final grades in math and language, and confidence in their own math abilities. The survey also elicited beliefs about the educational plans and college-going likelihood

[13] If they did not have an account, they were invited to create one and provide the handle.

[14] We conducted a baseline online survey of the head teachers, where we collected their phone numbers. This allowed us to later send WhatsApp videos to head teachers in the T2 schools.

of students' three closest friends, as well as measures of gender attitudes and self-efficacy. Importantly, participation required students to provide their national ID number, enabling us to track their post-secondary enrollment using administrative records.

The intervention was implemented over a six-week period in September and October 2023, during which nine videos were posted by the participating college students. Immediately after the final video was shared, in mid-October 2023, we invited all students in the study sample to complete a follow-up survey, using the same implementation strategy as at baseline, i.e., promotion by teachers in class and access via an online link. Approximately six weeks later, toward the end of the academic school year in December 2023, we announced a small scholarship program targeting students planning to pursue engineering in college. Specifically, we informed the sampled schools that two scholarships—one for a girl and one for a boy—would be randomly awarded among eligible applicants. To qualify, students were required to submit a brief application by early January 2024 and subsequently enroll in an engineering program in the following academic year. Importantly, in December 2024, we were able to obtain student's college enrollment data from the Education Ministry. Enrollment records, as of December 2024, i.e., approximately one year after high school graduation, were linked to each student via their unique national identification number, which we obtained during the survey data collection.

This data collection process led to our pre-registered outcomes of interest:

1. **Short-Term Outcome:** Self-stated *preferences for the engineering major* as generated by the follow-up survey;

2. **Medium-Term Outcome:** The decision to complete the *scholarship application* in December or January 2024. Given that the application took some time to fill in, and the selection was restricted to students enrolled in engineering in college, the application provides a revealed preference measure for students' intention to pursue engineering studies; it was reserved to ;

3. **Longer-Term Outcome:** *College Enrollment status*, as of December 2024, i.e., about one year after the high school graduation. While our primary outcome of interest is the decision to study engineering in college, this is conditional on college enrollment. Therefore, our longer-term primary outcomes are *college enrollment* and *engineering enrollment.*

Since all study participants were minors, both student assent and parental consent were obtained prior to participation. This led to a sample size of 547 girls, which correspond to about 15 percent of the graduating class in the sampled schools. When implementing the follow-up survey, only the 547 study participants were invited to fill it in. A total of 490 girls responded, i.e., about 90 percent of the original sample, with no significant differences in response rates across treatment arms.[15] For the scholarship application and college enrollment outcomes, however, we were able to obtain administrative data for the full sample of 547 students.

[15] We conduct and discuss attrition analysis in Section 4.1.

In addition to the girls' sample, we also collected survey data from a sample of 240 boys attending the same (co-ed) schools. Of these, 212 boys (88 percent) completed the follow-up survey. As with the girls, we obtained administrative data on scholarship applications (which were open to both boys and girls) and college enrollment for the full sample of 240 boys. We use the boys' sample both to document gender gaps in educational preferences within this context and as a placebo test, since the interventions were designed to target girls exclusively.

3.3 Estimation Strategy

Given the stratified school-level randomization and the individual-level nature of the outcomes of interest, we test for the impacts of the *Instagram (T1)* and the *Instagram & WhatsApp (T2)* intervention by estimating the following equation:

$$Y_{is} = \alpha + \beta_1 T1_s + \beta_2 T2_s + \gamma SES_s + \delta Coed_s + \lambda_s + \zeta \mathbf{X}_{is} + \epsilon_s \tag{1}$$

where Y_{is} represents the outcome for child i in school s; $T1_s$ and $T2_s$ indicate, respectively, the T1 and T2 treatment assignments for grade school s. We always control for the variables used to stratify the randomization: an indicator for the school being co-ed, an indicator for the school being categorized as low Socio-Economic Status (SES), and municipality fixed effects λ_s. In the most comprehensive specification, we control for child characteristics X and is selected for each outcome through a Double Lasso procedure (Belloni et al., 2014), out of a large set of covariates measured at baseline.[16] When the outcome of interest is a survey measure elicited at endline, X_{is} also includes the corresponding variable measured at baseline. Standard errors are clustered at the school level, our unit of randomization. Since we have four primary outcomes of interest, two obtained in the short- and medium-term through survey and scholarship application, and two obtained a year post graduation from administrative data on college enrollment, we correct for multiple hypothesis testing by employing the Romano-Wolf step-down procedure (Romano and Wolf, 2005) for each set of outcomes.

The main analysis is restricted to female students, since they were targeted by the Instagram intervention, and by the content of the WhatsApp videos sent to head teachers in the T2 schools. Since we have collected data also for a subset of boys in the sampled schools, we replicate the analysis for boys and for the combined sample, while including a gender dummy and its interaction with the treatment indicators. We report the resulting estimates in the Appendix.

Equation (1) allows us to estimate intent-to-treat effects. Since for every girl in T1 and T2 we know whether or not they followed the account, we are able to also obtain the Local Average Treatment Effects (LATE) by estimating a two-stage least squares regression. In the first stage, the decision to follow the

[16] The Double Lasso method performs two separate Lasso regressions, one regression of the outcome on all covariates to identify relevant predictors of the outcome, and another regression of the treatment on all covariates to identify predictors of the treatment. It then takes the union of the selected covariates from both regressions as the control vector.

account is predicted by the instrument, which is the initial assignment to treatment groups; in the second stage, the outcome of interest is regressed against the treatment take-up, as predicted in the first stage. Our primary regression tables report both ITT and LATE estimates.

We expect the impact of the interventions on interest in the economics major to vary by girls' math abilities. Therefore, following our pre-registered, we examining heterogeneous impacts by math aptitude by estimating the following equation:

$$\begin{aligned} Y_{is} = \alpha + \beta_1 T1_s + \beta_2 T2_s + \beta_3 T1_s * TopMath_i + \beta_4 T2_s * TopMath_i + \\ + \beta_5 TopMath_i + \gamma SES_s + \delta Coed_s + \lambda_s + \zeta \mathbf{X}_{is} + \epsilon_s \end{aligned} \tag{2}$$

where *TopMath* is a 0-1 indicator equal to 1 if the student received top grades in math at the end of the previous semester, i.e., a grade of 19 or 20, out of 20.[17] The estimated β_1 and β_2 in this case are the impacts of teh interventions on students who did not receive top grades in math, and the linear combinations of β_1 and β_3, and of β_2 and β_4 are the estimated impacts of T1 and T2, respectively, on top students.

4 Results

4.1 Descriptive Statistics

We present descriptive statistics of the schools involved in the study in Table 1 and of the sampled girls in Table 2. The vast majority of the schools are co-ed, with similar proportions across groups (84 percent of the Control and T1 schools, and 87 percent in T2), and no statistically significant differences. The groups are also balanced in terms of student-teacher ratios, which average around 17 across all arms, and in the proportion of students who are girls. Schools in the sample enroll between 400 and 500 students on average, with graduating cohorts ranging from 80 to 100 students. T1 schools tend to be larger than those in the Control and T2 groups, although the differences are not statistically significant. On average, each school has approximately 50 girls in their final year of high school, of whom between 13 and 20 percent participated in our study. In order to capture possible differences in the socio-economic backgrounds of the students served by the schools, we rely on the school classification system used by the Universidad de Piura in its admissions process, which allows us to distinguish between low- and high-SES schools. The majority of schools in the sample are classified as low SES. Although randomization was stratified by this variable, the small total number of schools led to some imbalance, with a higher share of low-SES schools in T1 (75 percent) and T2 (65 percent) compared to the Control group (56 percent). However, non-parametric tests indicate that these differences are not statistically significant.

Table 2 reports girls' characteristics, as recorded through our baseline survey. The table display a large number of variables, ordered by demographics (Panel A), self-reported confidence in math ability, completing

[17]This corresponds to the top quantile of the grade distribution.

college and studying engineering (Panel B), perceptions of parents and teachers' support (Panel C) and social media usage (Panel D). Overall, the sample appears well balanced across treatment arms, with few statistically significant differences. In terms of demographics (Panel A), girls are just over 16 years old on average, with a slightly higher mean age in T1 (16.22) compared to the Control group (16.08) and T2 (16.12). Less than one-third of girls have mothers or fathers who completed college, with the proportion being lower (277 and 26 percent for mothers and fathers respectively) in the T1 group, possibly reflecting the larger proportion of low SES schools in this group, observed in Table 1. Over 80 percent of girls aspire to a college education, but only 13 percent wish to study engineering in college.

Panel B of Table 2 shows overall balance across treatment groups in measures of girls' confidence in their academic abilities. Confidence in completing college is high across groups, as is confidence in math abilities, with one-fourth to one-third of girls believing they are among the top 20 percent of math performers in their class. Confidence in having the skills to pursue engineering, specifically industrial or mechanical engineering, is noticeably lower. On a scale from 1 (not confident at all) to 5 (extremely confident), average confidence levels are approximately 3 for industrial engineering and 2.6 for mechanical engineering, indicating a significant gap between general academic confidence and confidence in pursuing traditionally male-dominated STEM fields.

Panel C of Table 2 presents girls' perceptions of their parents' and teachers' aspirations for their future, as well as measures of gender attitudes and self-efficacy. Nearly all girls believe that their mothers and fathers want them to attend college, with averages above 90 percent across all groups. Perceived parental encouragement to pursue engineering is much lower, as only about 10 to 18 percent of girls believe their mother or father wishes them to study engineerin, with no significant differences across treatment arms. However, perceptions of teacher support show some imbalance: a significantly higher proportion of girls in the treatment groups (74 percent in T1 and 70 percent in T2) report that they discuss possible college enrollment with their teachers, compared to 58 percent in the control group. In contrast, few girls report that their teacher recommended engineering as field of study to them (under 12 percent in all groups), with no statistically significant differences. Finally, measures of gender attitudes and self-efficacy are balanced across groups, with average values close to zero due to standardization of these measures around the Control group's mean.

Panel D of Table 2 indicate that nearly 90 percent of girls had an Instagram account at baseline, and almost 50 percent spent over 2 hours daily on social media. Finally, we see that about 30 percent of girls in the treatment schools followed the Instagram account we created, as opposed to none of the girls in the Control schools.

While our study targeted girls, we also collecting survey data from a sample of final year boys attending the same (co-ed) high schools. In Table A1, in Appendix, we compare the average characteristics of the surveyed girls and boys. We observe significant gender gaps across several dimensions. While academic performance in math and language is similar across genders, boys report substantially higher levels of self-confidence. Nearly 58 percent of boys believe they are among the top 50 percent of math performers in their class, compared to just 29 percent of girls. Similarly, a greater share of boys (42 percent) believe they are in the top 20 percent, compared to 24 percent of girls. Confidence in completing college is high for both genders but marginally higher among boys. The gender gap widens further when examining confidence in the ability to pursue engineering: boys report significantly higher confidence in their ability to study both industrial and mechanical engineering, with average scores of 3.65 and 3.44, respectively, compared to 3.00 and 2.63 among girls. Despite these gaps in self-confidence, girls report slightly higher college aspirations, with 87 percent wishing to attend college compared to 84 percent of boys.

Panel C of Table A1 highlights that students perceive markedly different levels of parental and teacher encouragement for pursuing engineering, depending on gender. While almost all students report that their parents want them to attend college, boys are much more likely to believe their parents, and especially mothers, wish for them to pursue engineering. A similar pattern emerges in perceptions of teachers' recommended fields of study: 22 to 27 percent of boys mention engineering as the field of study most often recommended to them by their teachers, as opposed to only around 10 percent of girls. These differences suggest that boys may receive stronger signals of support for entering traditionally male-dominated STEM fields, both at home and in the classroom.[18] Finally, boys are less likely to have an Instagram account and spend less time than girls on social media.

4.2 Short- and Medium-Term Treatment Impacts on Preferences

In this section, we present findings on the impact of the social media interventions on girls' preferences for pursuing an engineering major in college. The endline survey was conducted online immediately following the completion of the intervention. Only students who had participated in the baseline survey and for whom we had obtained parental consent were invited to take the follow-up survey. Approximately 6 to 8 weeks later, we invited all students in the sampled schools to apply for a scholarship to study engineering in college, as described in Section 3.2. About 90 percent (N=490) of the initially surveyed girls participated in the follow-up survey. We do not see any differential attrition by treatment status. We also do not find any significant predictors of attrition among individual and school characteristics, as shown in Table A2 in Appendix.

Table 3 reports the estimated intent-to-treat effects (Panel A) and local average treatment effects (Panel B) of the social media interventions on girls' self-reported interest in engineering and on actual scholarship application behavior. In Table 4, we present estimates from a pre-registered heterogeneity analysis based on academic performance in math. Specifically, we compare the impacts of the intervention between top-in-math girls and the rest of the girls.[19] Figure 2 displays the estimated coefficients generated for T1 and T2 for the full sample (from Table 3), and for girls who are either top or not top performers in math (from

[18]We also see that boys exhibit more regressive gender attitudes on average (with more negative standardized scores) and significantly higher levels of self-efficacy.

[19]The measure of math performance is based on self-reported final year grade in math.

Table 4).

Our estimates indicate that the social media interventions had an immediate negative effect on girls' self-reported interest in pursuing an engineering major. This effect is statistically significant among girls who are not top performers in math, particularly when their teachers were also targeted by the intervention. Among top-performing girls, the effects are noisier. However, our revealed measure of preferences for engineering – the scholarship applications – reveals a different pattern. While we find no overall effect of the interventions on applications, the impact is positive, large, and statistically significant among top-performing girls, particularly in T2, when the head teachers also received videos from professors encouraging them to promote the engineering major among their female students.

4.3 Longer-Term Treatment Impacts on College Enrollment

We report the estimated treatment effects on college enrollment outcomes in Tables 5 and 6. The estimated coefficients for the full sample, the top-in-math and the not-top girls are displayed in Figure 3. Panel A of Table 5 presents Intent-to-Treat (ITT) estimates, while Panel B shows Local Average Treatment Effects (LATE). Columns 1 and 2 focus on enrollment in engineering, and columns 3 and 4 on overall college enrollment, with and without Double LASSO controls. The results indicate no significant effects of either treatment on engineering enrollment across all specifications (columns 1–2). However, both treatments had a significant positive impact on overall college enrollment (columns 3–4). Specifically, in the ITT estimates, T1 led to a 20 percentage point increase in college enrollment (corresponding to a 70 percent increase over the control mean), while T2 had a smaller but still significant effect of 10 percentage points (a 38 percent increase). The LATE estimates in Panel B reinforce these findings, showing even larger treatment effects—T1 increased enrollment by 70 percentage points, and T2 by 35 percentage points. While the Instagram intervention (T1) appears more effective when implemented alone than when combined with the teacher intervention (T1), the differences in impacts are not statistically significant. The large treatment effects on college enrollment are robust to the inclusion of LASSO controls (column 4) and to correcting p-values for multiple hypothesis testing. When replicating the analysis for boys, who were not targeted by the intervention, we see, as expected, null impacts on all variables of interest, as shown in Table A3 in Appendix.

In Table 6, we estimate equation 2 of Section 3.3 to examine whether the social media interventions had differential impacts based on students' math abilities. We find no effects of either intervention on enrollment in engineering programs, regardless of girls' math aptitude. In contrast, we observe notable heterogeneity in the effects on overall college enrollment. For this outcome, the Instagram (T1) and Instagram & WhatsApp (T2) interventions show distinct patterns depending on students' math ability. Specifically, the Instagram intervention increases enrollment among non-top-in-math girls by 20 percentage points, but this effect disappears when teachers are also targeted in the T2 treatment. For T2, we find no statistically significant impact on non-top girls. Conversely, the Instagram intervention alone (T1) is not sufficient to increase enrollment among top-in-math girls. For this group, it is the combination of the Instagram intervention and the WhatsApp-based engagement with head teachers that leads to a significant increase in college enrollment. These findings suggest that teachers play an important role in supporting girls' post-secondary transitions. However, under T2, where teachers were encouraged to advise high achieving girls to pursue engineering studies, they may have reallocated their advising and encouragement efforts more heavily toward top students, thereby diminishing the support available to non-top girls.

4.4 Analysis of Mechanisms

We pre-registered two sets of secondary outcomes to help uncover the mechanisms driving the observed treatment effects. Specifically, we planned to examine the impact of the social media interventions on: (1) students' confidence in their math ability; and (2) their perceptions of parental and teacher preferences and support regarding post-secondary education. In this section, we present the findings from this secondary analysis. To measure students' beliefs about their parents' educational preferences, we asked whether their parents wanted them to attend college after high school and, if so, which major they preferred. To assess perceived support from teachers, we asked students whether they had discussed college enrollment with a teacher, and if so, what major the teacher (or teachers) had recommended the most. We are particularly interested in whether teachers encouraged students to pursue engineering. In addition, we thought it was important to measure students' beliefs about their best friends' post-secondary educational choices. To do so, we asked each student to name their three closest friends and, for each one, report whether they believed that friend would attend college and, if so, what major they would pursue. Our key outcomes in this domain are: (1) the number of friends (out of three) whom the respondent believes will go to college, and (2) among those friends expected to attend college, the number believed to be pursuing an engineering degree.[20]

Regarding students' confidence, we had originally intended to test whether the interventions made girls feel more confident in their math ability. However, when finalizing the survey, we realized that a most plausible factor activated by the intervention would be confidence in successfully completing college, and confidence in successfully obtaining an engineering degree. We therefore asked students for their level of confidence in their ability to pursue a college degree, as well as confidence in their ability to complete specific degrees: i) a mechanical engineering degree; ii) an industrial engineering degree; iii) a law degree; iv) an economics or business degree; v) a medicine degree; vi) an education degree.[21]

Table 7 presents the estimated impacts of the two social media interventions on self-reported confidence in students' ability to complete college and to complete an engineering degrees. Confidence is measured in two ways: a 5-item Likert confidence score, where 1 indicates a "definitely no" answer and 5 indicates a "definitely yes" answer to the question: " Do you believe that you have the academic aptitude and abilities to complete a college degree/a college degree in mechanical engineering/ a college degree in industrial engineering?"

[20]These measures are not pre-registered.

[21]Due to a programming mistake, we failed we include in the follow-up survey the question we had included at baseline to record confidence in own math abilities.

(literally translated from Spanish). For ease of interpretation, We have standardized this variable around the control mean.

While we find no significant impacts on students' confidence in their ability to complete a mechanical or industrial engineering degree,[22] we observe a significant increase in confidence to complete college among students exposed to the Instagram intervention (T1). This effect disappears when teachers are also targeted through the WhatsApp intervention (T2). Our heterogeneity analysis further reveals that these differences between the impacts of T1 and T2 is driven by non-top students, whose confidence in completing college increases only under T1. In fact, the point estimate for T2 is slightly negative for this group, although not statistically significant. In contrast, top students experience a boost in college-related confidence under both interventions, suggesting that teacher targeting in T2 may have led to a reallocation of support and encouragement disproportionately toward high-achieving students. This pattern reinforces the interpretation that confidence is a key mechanism linking the Instagram intervention to increased college enrollment for non-top girls, and that the targeting of teachers in T2 may have unintentionally disrupted this channel by undermining the confidence of lower-performing students.

Table 8 examines whether the social media interventions influenced students' perceptions of their parents', teachers', and peers' preferences and support regarding post-secondary education. The results show no significant changes in girls' beliefs about their parents' preferences, either regarding college attendance or the pursuit of engineering. Likewise, students' perceptions of teacher's support remain unaffected, as we find no impact of the interventions on the likelihood that students discussed college with a teacher or received encouragement to pursue engineering. In contrast, we find significant and positive treatment effects on students' perceptions of their peers' educational aspirations. Specifically, students exposed to either intervention were more likely to believe that their close friends would go to college. Both T1 and T2 increased the number of friends a student believes will attend college. However, neither intervention increased the number of friends believed to pursue engineering; if anything the impact on these beliefs is negative, mirroring the negative immediate impacts of T1 and T2 on preferences for the engineering major observed in Table 3. Taken together, these results suggest that the impacts of the social media interventions on college enrollment may have been amplified by peer effects—specifically, by shifts in girls' beliefs about their close friends' likelihood of attending college. This is especially true for not top-in-math, as shown in Figure 5 and in Table A6 in Appendix.

5 Discussion and Conclusion

This paper investigates the potential of social media to influence girls' post-secondary educational decisions in a relatively low-income setting. Through a randomized controlled trial in Northern Peru, we evaluated

[22] We also do not find robust evidence of impacts of completing other college degrees, with the exception of some weak impacts of T2 on confidence in pursuing a medicine degree and an economics degrees. These estimates are displayed in Table A4 in Appendix.

the impact of an information and role model campaign delivered via Instagram. In one treatment arm, we complemented the Instagram intervention with an information campaign targeting head teachers via WhatsApp. The primary aim was to inspire more girls in their final year of high school to consider and pursue STEM fields, and engineering in particular. Our results highlight both the promise and complexity of using social media to shift educational trajectories.

First, we find that the Instagram intervention alone had a large and significant effect on overall college enrollment. Girls exposed to T1 were 20 percentage points more likely to be enrolled in college one year after graduating high school, representing a 75 percent increase over the control mean of 26.5 percent. When the WhatsApp teacher component was added, we saw a positive but smaller effect. However, neither treatment significantly increased enrollment in engineering programs. In the short term, the treatments actually reduced self-reported interest in engineering, particularly among girls who were not top performers in math. These students may have felt discouraged by the demanding academic imagery portrayed in the videos or by heightened awareness of the challenges associated with engineering studies. Despite this short-term decline in interest, we observe that the interventions positively affected revealed preferences among high-achieving girls. Top-in-math students exposed to both Instagram content and teacher encouragement were significantly more likely to apply for a scholarship to study engineering, suggesting that the combination of peer role models and teacher reinforcement may help activate latent interest in STEM fields. However, this did not translate into higher enrollment rates in engineering a year after the high school graduation.

Our analysis of mechanisms suggests that confidence in the ability to succeed in higher education, rather than in specific fields like engineering, played a central role in driving the increase in college enrollment. Exposure to the Instagram content increased confidence significantly under T1, especially among non-top-in-math girls, but did not when the head teachers were also targeted, in T2. This pattern aligns with the heterogeneity in enrollment impacts and suggests that the inclusion of teachers may have shifted their attention and encouragement disproportionately toward top-performing students, inadvertently crowding out lower-achieving girls. A second mechanism behind the success of the intervention appears to be peer influence. Exposure to the intervention increased girls' beliefs that their close friends would attend college. These peer-related changes likely reinforced the enrollment effects of the interventions.

The large treatment effects we observe on college enrollment, especially under the Instagram-only intervention, are striking, and likely reflect not only the direct influence of the role model content but also important amplification mechanisms. Beyond peer effects operating through changes in beliefs about close friends' educational trajectories, the Instagram algorithm itself may have reinforced the intervention's impact. Once students began engaging with the curated role model content, Instagram's recommendation system may have exposed them to additional, similar content, such as videos and accounts created by other young women documenting their college experiences. This algorithmic spillover could have broadened and deepened the perceived social norm around college attendance, increasing its salience and perceived attainability. In this way, the intervention may have created a positive feedback loop between students' own aspirations, peer dynamics, and the broader digital environment they navigate daily.

Overall, our study highlights the promise of social media as a tool for expanding educational opportunities for teenagers. Our findings suggest that reaching and inspiring teenage girls through digital platforms may be easier and more impactful than previously thought. As internet access and smartphone penetration continue to grow in low-income settings, social media platforms offer a scalable and cost-effective way of reaching adolescents, particularly those who might otherwise lack access to role models, family support, or institutional guidance.

References

Agurto, M., M. Bazan, S. Hari, and S. Sarangi (2023). To inspire and to inform: The role of role models. *Available at SSRN 4595079*.

Alan, S., S. Ertac, and I. Mumcu (2018). Gender stereotypes in the classroom and effects on achievement. *Review of Economics and Statistics 100*(5), 876–890.

Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020). The welfare effects of social media. *American economic review 110*(3), 629–676.

Anderson, M., M. Faverio, and J. Gottfried (2023). Teens, social media and technology 2023. *Pew Res Cent*.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (2024a). The economics of social media. *Journal of Economic Literature 62*(4), 1422–1474.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (2024b). Experiments on social media. Technical report, CESifo.

Barr, A. C. and B. L. Castleman (2025). Increasing degree attainment among low-income students: The role of intensive advising and college quality. Technical report, National Bureau of Economic Research.

Beland, L.-P. and R. Murphy (2016). Ill communication: technology, distraction & student performance. *Labour Economics 41*, 61–76.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies 81*(2), 608–650.

Bettinger, E. P. and B. J. Evans (2019). College guidance for all: A randomized experiment in pre-college advising. *Journal of Policy Analysis and Management 38*(3), 579–599.

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.

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

Bursztyn, L., B. R. Handel, R. Jimenez, and C. Roth (2023). When product markets become collective traps: The case of social media. Technical report, National Bureau of Economic Research.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Carrell, S. and B. Sacerdote (2017). Why do college-going interventions work? *American Economic Journal: Applied Economics 9*(3), 124–151.

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.

Di, A. (2024). Inspirational or demotivational: Experimental evidence on the impacts of role models. *Available at SSRN 4749305*.

Kessel, D., H. L. Hardardottir, and B. Tyrefors (2020). The impact of banning mobile phones in swedish secondary schools. *Economics of Education Review 77*, 102009.

Kipchumba, E., C. Porter, D. Serra, and M. Sulaiman (2024). The impact of role models on youths' aspirations, gender attitudes and education in somalia.

Kreski, N., J. Platt, C. Rutherford, M. Olfson, C. Odgers, J. Schulenberg, and K. M. Keyes (2021). Social media use and depressive symptoms among united states adolescents. *Journal of Adolescent Health 68*(3), 572–579.

Lavy, V. and R. Megalokonomou (2024). The short-and the long-run impact of gender-biased teachers. *American Economic Journal: Applied Economics 16*(2), 176–218.

Lim, J. and J. Meer (2020). Persistent effects of teacher–student gender matches. *Journal of Human Resources 55*(3), 809–835.

More Sánchez, J. and D. Argandoña Martinez (2020). Las redes de transporte de fibra óptica, microondas y satelital y su rol para promover la expansión de la cobertura de los servicios públicos de telecomunicaciones: reporte y mapas de cobertura. documento de trabajo nº 49.

Mosquera, R., M. Odunowo, T. McNamara, X. Guo, and R. Petrie (2020). The economic effects of facebook. *Experimental Economics 23*(2), 575–602.

Porter, C. and D. Serra (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics 12*(3), 226–254.

Romano, J. P. and M. Wolf (2005). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Serra, D. (2025). Role models in developing countries. In *Handbook of experimental development economics*, pp. 123–135. Edward Elgar Publishing.

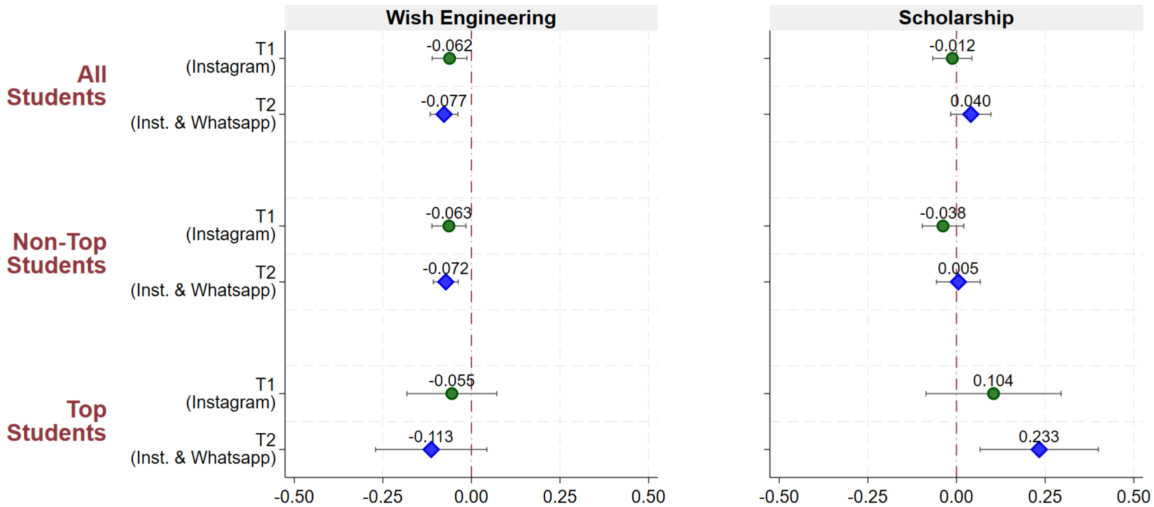
Sevilla, M. P., P. Bordón, and F. Ramirez-Espinoza (2023). Reinforcing the stem pipeline in vocational-technical high schools: The effect of female teachers. *Economics of Education Review 95*, 102428.

Figures and Tables

Figure 1: The Instagram Account

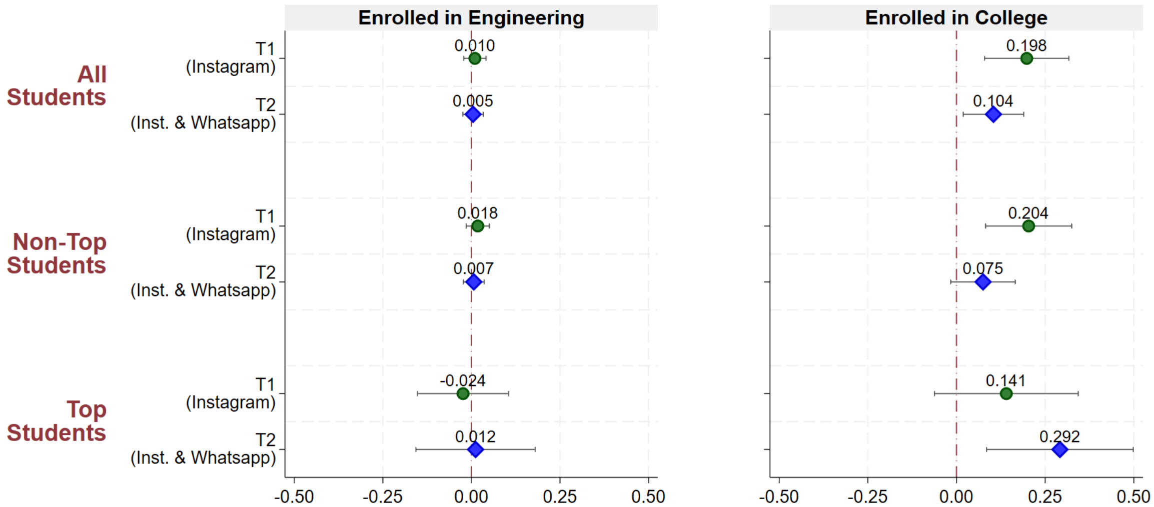
Note: The figure displays the Instagram account logo and name, as well as the first post (on the left), which states: " We are a group of university students and we are excited to share our life in university with you. Stay tuned to learn about our experiences, academic advice and special moments that make this stage unforgettable."

Figure 2: Coefficient Plots of Estimated Short-Term and Medium-Term Impacts



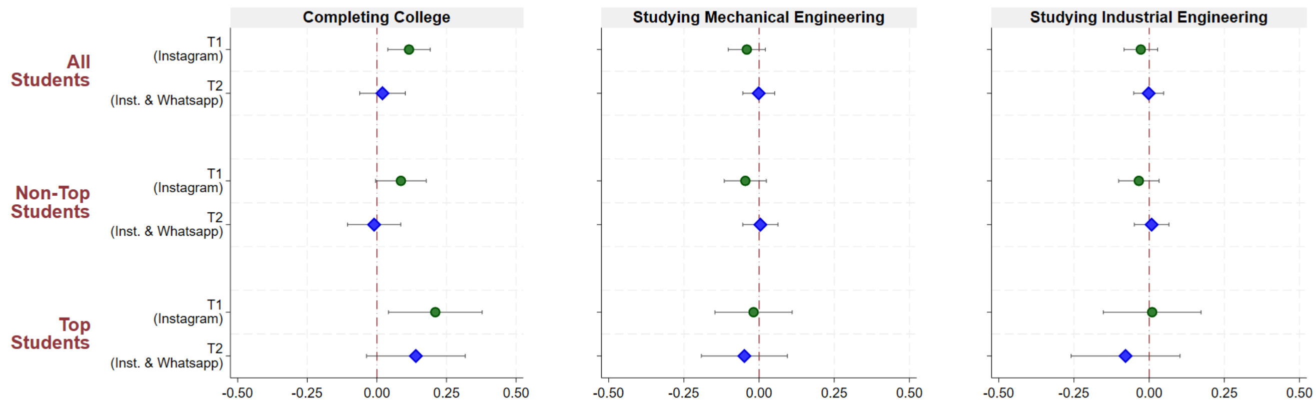
Notes: Coefficient estimates come from ITT regressions reported in Tables 3 and 4. Controls are selected through the Double Lasso procedure from the set of variables displayed in Table 2, and always included a school-level indicators for low SES status, and co-educational status, in addition to metropolitan area fixed effects. The error bars correspond to 95% confidence intervals.

Figure 3: Coefficient Plots of Longer-Term Impacts on Enrollment Outcomes



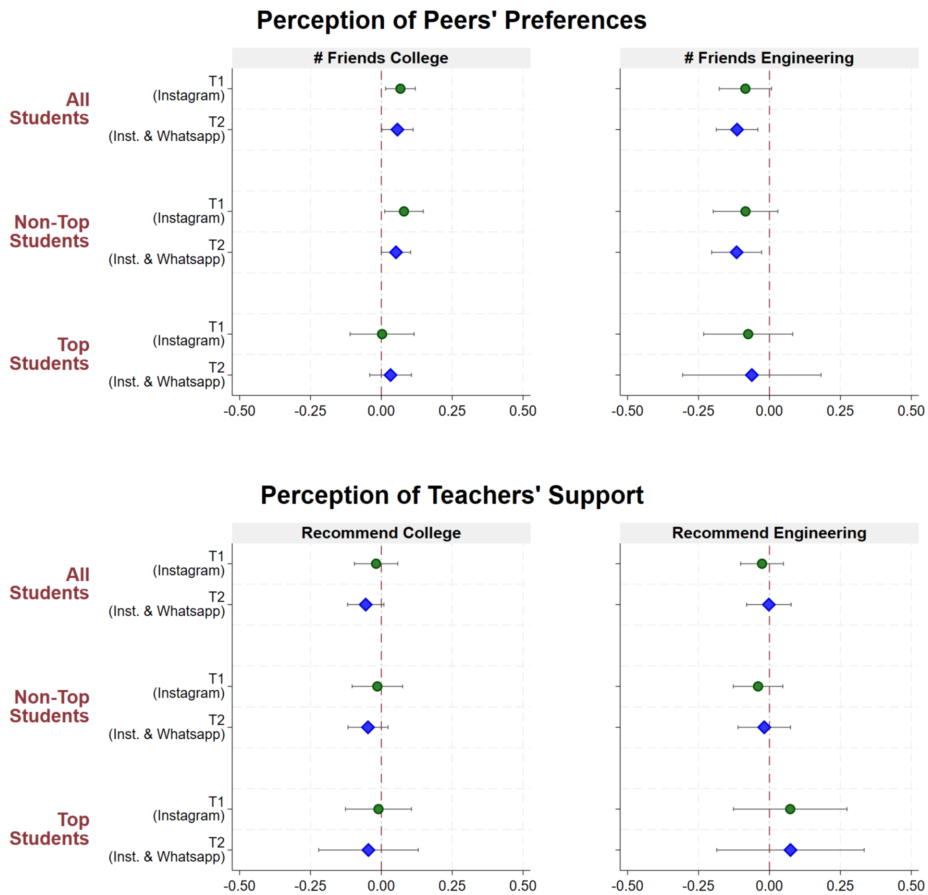
*Notes:*Coefficient estimates come from ITT regressions reported in Tables 5 and 6. Controls are selected through the Double Lasso procedure from the set of variables displayed in Table 2, and always included a school-level indicators for low SES status, and co-educational status, in addition to metropolitan area fixed effects. The error bars correspond to 95% confidence intervals.

Figure 4: Self-reported Confidence



*Notes:*Coefficient estimates come from ITT regressions reported in Tables 7 and A5. Controls are selected through the Double Lasso procedure from the set of variables displayed in Table 2, and always include a school-level indicators for low SES status, and co-educational status, in addition to metropolitan area fixed effects. The error bars correspond to 95% confidence intervals.

Figure 5: Perception of Peers' Preferences and Teachers' Support



Notes: Coefficient estimates come from ITT regressions reported in Tables 8 and A6. Controls are selected through the Double Lasso procedure from the set of variables displayed in Table 2, and always include a school-level indicators for low SES status, and co-educational status, in addition to metropolitan area fixed effects. The error bars correspond to 95% confidence intervals.

Table 1: School Balance Table

	Control	T1	T2	C=T1	C=T2	T1=T2	C=T1=T2
Co-ed	0.84	0.84	0.87	1.000	0.774	0.774	0.950
	(0.37)	(0.37)	(0.34)				
Low SES	0.56	0.76	0.65	0.232	0.519	0.417	0.339
	(0.51)	(0.44)	(0.49)				
Total Students	388.92	513.08	405.87	0.241	0.710	0.599	0.502
	(374.58)	(496.96)	(301.54)				
Student/Teacher Ratio	17.93	17.02	17.45	0.519	0.489	0.898	0.930
	(8.49)	(7.92)	(8.66)				
Perc. Female Teachers	0.47	0.39	0.38	0.353	0.201	0.915	0.250
	(0.23)	(0.19)	(0.14)				
Total Final-Year Boys	28.00	46.72	37.78	0.420	0.577	0.901	0.453
	(24.11)	(77.31)	(39.35)				
Total Final-Year Girls	50.12	54.76	41.96	0.444	0.718	0.489	0.738
	(76.17)	(51.33)	(35.24)				
Perc. Final-Year Girls	0.54	0.54	0.54	0.776	0.772	0.967	0.992
	(0.24)	(0.21)	(0.21)				
Average Number of Boys Surveyed	3.08	2.56	4.30	0.796	0.458	0.327	0.206
	(3.28)	(2.40)	(4.42)				
Average Number of Girls Surveyed	6.48	7.40	8.70	0.805	0.258	0.449	0.617
	(6.60)	(8.44)	(8.26)				
Observations	25	25	23				

Notes: The table reports average school characteristics, by treatment assignment. For each variable, the p-values reported in columns 5 to 7 are generated by non-parametric Wilcoxon rank-sum tests of equality between the groups.

Table 2: Balance Table - Girls

	Control	T1	T2	C=T1	C=T2	T1=T2	C=T1=T2
Panel A: Demographics							
Age	16.08	16.22	16.12	0.006	0.292	0.026	0.008
	(0.46)	(0.45)	(0.35)				
Mom Went to College	0.32	0.27	0.31	0.236	0.847	0.292	0.439
	(0.47)	(0.45)	(0.47)				
Dad Went to College	0.34	0.26	0.32	0.048	0.604	0.119	0.121
	(0.47)	(0.44)	(0.47)				
Math Performance	16.56	16.73	16.34	0.437	0.328	0.052	0.172
	(2.21)	(1.96)	(1.97)				
Language Performance	17.33	17.30	17.23	0.847	0.570	0.703	0.848
	(1.78)	(1.70)	(1.51)				
Wish College	0.87	0.87	0.83	0.998	0.289	0.271	0.436
	(0.34)	(0.34)	(0.38)				
Wish Engineering	0.14	0.13	0.13	0.740	0.741	0.994	0.931
	(0.35)	(0.34)	(0.34)				
Wish STEM	0.17	0.15	0.14	0.589	0.392	0.753	0.687
	(0.38)	(0.36)	(0.35)				
Panel B: Self-Reported Confidence in...							
Being in Top 50% (math)	0.32	0.29	0.29	0.487	0.595	0.855	0.770
	(0.47)	(0.45)	(0.46)				
Being in Top 20% (math)	0.33	0.24	0.24	0.065	0.066	0.960	0.103
	(0.47)	(0.43)	(0.43)				
Completing College	4.38	4.35	4.38	0.756	0.947	0.694	0.918
	(0.73)	(0.72)	(0.74)				
Studying Industrial Engineering	3.06	3.00	2.85	0.681	0.115	0.260	0.275
	(1.19)	(1.31)	(1.21)				
Studying Mechanical Engineering	2.72	2.63	2.54	0.505	0.156	0.457	0.371
	(1.23)	(1.25)	(1.18)				
Panel C: Perceptions							
Mom Wishes College	0.98	0.94	0.98	0.053	0.919	0.046	0.045
	(0.14)	(0.24)	(0.14)				
Dad Wishes College	0.98	0.91	0.95	0.009	0.303	0.067	0.017
	(0.16)	(0.29)	(0.21)				
Mom Wishes Engineering	0.14	0.10	0.12	0.264	0.637	0.493	0.535
	(0.35)	(0.30)	(0.33)				
Dad Wishes Engineering	0.18	0.17	0.17	0.883	0.726	0.835	0.939
	(0.38)	(0.38)	(0.37)				
Teacher Wishes College	0.58	0.74	0.70	0.002	0.018	0.446	0.006
	(0.50)	(0.44)	(0.46)				
Teacher Wishes Engineering	0.08	0.12	0.10	0.234	0.624	0.448	0.474
	(0.27)	(0.32)	(0.29)				
Gender Attitudes	-0.00	0.09	0.08	0.260	0.270	0.912	0.446
	(0.68)	(0.73)	(0.66)				
Self-Efficacy	-0.00	0.03	0.00	0.248	0.854	0.301	0.444
	(0.25)	(0.25)	(0.25)				
Panel D: Social Media							
Has IG Account	0.87	0.85	0.90	0.563	0.468	0.174	0.397
	(0.34)	(0.36)	(0.31)				
Spends >2hr on social media	0.47	0.46	0.49	0.857	0.694	0.550	0.829
	(0.50)	(0.50)	(0.50)				
Followed Account	0.00	0.27	0.32	0.000	0.000	0.287	0.000
	(0.00)	(0.45)	(0.47)				
Observations	162	185	200				

Notes: The table reports the characteristics of the surveyed girls, by treatment assignment. For each variable, the p-values reported in columns 5 to 7 are generated by tests of equality of means. All variables are generated by the baseline survey students participated in, in July 2023. The math and language performance averages in Panel A refer to the final grade obtained at the end of the previous academic year. The highest possible grade is 20. In panel P, confidence in being a top student indicates the percentage of students who answered "yes, definitely" to the respective question. The other confidence questions are average answers on a scale from 1 ("definitely no") and 5 ("definitely yes"). The Gender Attitudes and Self-Efficacy measures are aggregated indexes, which have been standardized around the control mean.

Table 3: Short- and Medium-Term Impacts on Preferences and Scholarship Applications

	Wish Engineering		Scholarship	
	(1)	(2)	(3)	(4)
Panel A: ITT estimates				
Instagram (T1)	-0.067**	-0.062**	-0.016	-0.012
	(0.031)	(0.029)	(0.035)	(0.033)
	[0.017]	[0.027]	[0.533]	[0.655]
Instagram & WhatsApp (T2)	-0.086***	-0.077***	0.033	0.040
	(0.024)	(0.024)	(0.035)	(0.034)
	[0.004]	[0.004]	[0.235]	[0.154]
Observations	490	490	547	547
Clusters	67	67	73	73
Control Mean	0.185	0.185	0.111	0.111
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	–	–
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.415	0.496	0.157	0.147
Panel B: LATE estimates				
Instagram (T1)	-0.203**	-0.188**	-0.060	-0.044
	(0.099)	(0.092)	(0.117)	(0.111)
	[0.029]	[0.025]	[0.509]	[0.631]
Instagram & WhatsApp (T2)	-0.257***	-0.232***	0.097	0.121
	(0.072)	(0.069)	(0.105)	(0.102)
	[0.004]	[0.003]	[0.242]	[0.161]
Observations	490	490	547	547
Clusters	67	67	73	73
Control Mean	0.185	0.185	0.111	0.111
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	–	–
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.450	0.529	0.149	0.141

Notes: Panel A displays ITT estimates from linear probability models, where the dependent variables are 0-1 indicators for the student stating that engineering is their desired field of study in college (columns 1 and 2), and for the student applying for the engineering scholarship that we created (columns 3 and 4. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. When available, we include the outcome variable measured at baseline. Panel B displays IV regression estimates, where the instrument is the random assignment of schools to treatment or control groups. Controls in columns 2 an 4 are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses. Multiple hypothesis corrected p-values using the Romano-Wolf (2005) procedure are reported in square brackets.

Table 4: Impacts on Preferences and Scholarship

Heterogeneity by Top Performance in Math

	Wish Engineering		Scholarship	
	(1)	(2)	(3)	(4)
Instagram (T1)	-0.069**	-0.063**	-0.039	-0.038
	(0.031)	(0.029)	(0.036)	(0.035)
	[0.032]	[0.030]	[0.096]	[0.199]
Instagram & WhatsApp (T2)	-0.078***	-0.072***	0.001	0.005
	(0.021)	(0.021)	(0.038)	(0.037)
	[0.006]	[0.007]	[0.891]	[0.854]
Top	0.055	0.030	-0.083	-0.135
	(0.068)	(0.075)	(0.051)	(0.081)
T1 × Top	0.031	0.008	0.160	0.142
	(0.083)	(0.079)	(0.120)	(0.122)
	[0.599]	[0.888]	[0.058]	[0.214]
T2 × Top	-0.036	-0.041	0.233**	0.228**
	(0.096)	(0.096)	(0.095)	(0.107)
	[0.630]	[0.570]	[0.009]	[0.036]
Observations	490	490	547	547
Clusters	67	67	73	73
Control Mean	0.162	0.162	0.121	0.121
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	–	–
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.726	0.723	0.227	0.237
$H_0 : T1 + T1 \times Top = 0$	0.617	0.471	0.304	0.364
$H_0 : T2 + T2 \times Top = 0$	0.225	0.233	0.014	0.023
$H_0 : T1 + T1 \times Top = T2 + T2 \times Top$	0.363	0.469	0.387	0.311

Notes: The table displays ITT estimates from linear probability models, where the dependent variables are 0-1 indicators for the student stating that engineering is their desired field of study in college (columns 1 and 2), and for the student applying for the engineering scholarship that we created (columns 3 and 4. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. When available, we include the outcome variable measured at baseline. "Top" indicates whether the student was among the top performers (top quantile) in math, based on their final grade on the year preceding the study, i.e, the year before high school graduation. Controls in columns 2 an 4 are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses. Multiple hypothesis corrected p-values using the Romano-Wolf (2005) procedure are reported in square brackets.

Table 5: Longer-Term Impacts: College Enrollment (Administrative Data)

	Enrolled in Engineering		Enrolled in College	
	(1)	(2)	(3)	(4)
Panel A: ITT estimates				
Instagram (T1)	0.004	0.010	0.210**	0.198***
	(0.025)	(0.019)	(0.085)	(0.071)
	[0.861]	[0.558]	[0.003]	[0.004]
Instagram & WhatsApp (T2)	0.004	0.005	0.121*	0.104**
	(0.019)	(0.017)	(0.061)	(0.051)
	[0.869]	[0.779]	[0.049]	[0.069]
Observations	547	547	547	547
Clusters	73	73	73	73
Control Mean	0.068	0.068	0.265	0.265
Randomization Controls	Yes	Yes	Yes	Yes
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.990	0.813	0.252	0.164
Panel B: LATE estimates				
Instagram (T1)	0.013	0.035	0.718**	0.699**
	(0.086)	(0.066)	(0.342)	(0.304)
	[0.864]	[0.561]	[0.006]	[0.015]
Instagram & WhatsApp (T2)	0.011	0.016	0.387*	0.350**
	(0.058)	(0.052)	(0.203)	(0.178)
	[0.864]	[0.770]	[0.057]	[0.068]
Observations	547	547	547	547
Clusters	73	73	73	73
Control Mean	0.068	0.068	0.265	0.265
Randomization Controls	Yes	Yes	Yes	Yes
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.980	0.790	0.273	0.193

Notes: Panel A displays ITT estimates from linear probability models, where the dependent variables are 0-1 indicators for the student being enrolled in engineering as of December 2024 (columns 1 and 2), and for the student student being enrolled in college, regardless of major, as of December 2024 (columns 3 and 4. Panel B displays IV regression estimates. The instrument is the random assignment of schools to treatment or control groups. "Top" indicates whether the student was among the top performers (top quantile) in math, based on their final grade on the year preceding the study, i.e, the year before high school graduation. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. Controls in columns 2 an 4 are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses. Multiple hypothesis corrected p-values using the Romano-Wolf (2005) procedure are reported in square brackets.

Table 6: Longer-Term Impacts on Enrollment (Admin Data)

Heterogeneity by Top Performance in Math

	Enrolled in Engineering		Enrolled in College	
	(1)	(2)	(3)	(4)
Instagram (T1)	0.023	0.018	0.230**	0.204***
	(0.021)	(0.020)	(0.086)	(0.073)
	[0.483]	[0.297]	[0.001]	[0.004]
Instagram & WhatsApp (T2)	0.012	0.007	0.109	0.075
	(0.017)	(0.018)	(0.066)	(0.055)
	[0.716]	[0.682]	[0.083]	[0.219]
Top	0.112*	0.052	0.090	-0.016
	(0.066)	(0.072)	(0.080)	(0.097)
T1 × Top	-0.043	-0.042	-0.079	-0.063
	(0.083)	(0.083)	(0.110)	(0.114)
	[0.837]	[0.730]	[0.639]	[0.730]
T2 × Top	0.005	0.005	0.176	0.217
	(0.108)	(0.107)	(0.140)	(0.135)
	[0.578]	[0.953]	[0.161]	[0.102]
Observations	547	547	547	547
Clusters	73	73	73	73
Control Mean	0.045	0.045	0.258	0.258
Randomization Controls	Yes	Yes	Yes	Yes
LASSO Controls	No	Yes	No	Yes
$H_0 : T1 = T2$	0.609	0.578	0.128	0.059
$H_0 : T1 + T1 \times Top = 0$	0.796	0.761	0.232	0.252
$H_0 : T2 + T2 \times Top = 0$	0.861	0.908	0.034	0.022
$H_0 : T1 + T1 \times Top = T2 + T2 \times Top$	0.691	0.706	0.328	0.281

Notes: The table displays ITT estimates from linear probability models, where the dependent variables are 0-1 indicators for the student being enrolled in engineering as of December 2024 (columns 1 and 2), and for the student student being enrolled in college, regardless of major, as of December 2024 (columns 3 and 4). "Top" indicates whether the student was among the top performers (top quantile) in math, based on their final grade on the year preceding the study, i.e, the year before high school graduation. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. Controls in columns 2 an 4 are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses. Multiple hypothesis corrected p-values using the Romano-Wolf (2005) procedure are reported in square brackets.

Table 7: Mechanisms: Impacts on Self-Reported Confidence

	Completing College		Industrial Engineering		Mechanical Engineering	
	Confidence	Def. yes	Confidence	Def. yes	Confidence	Def. yes
	(1)	(2)	(3)	(4)	(5)	(6)
Instagram (T1)	0.154	0.115**	-0.044	-0.028	-0.091	-0.041
	(0.093)	(0.046)	(0.059)	(0.033)	(0.077)	(0.037)
Instagram & WhatsApp (T2)	0.077	0.020	-0.029	-0.002	-0.062	-0.001
	(0.084)	(0.049)	(0.074)	(0.030)	(0.086)	(0.032)
Observations	463	463	490	490	490	490
Control Mean	0.000	0.579	-0.000	0.178	0.000	0.164
Randomization Controls	Yes	Yes	Yes	Yes	Yes	Yes
LASSO Controls	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.436	0.041	0.820	0.368	0.709	0.134

Notes: The table displays ITT estimates generated by OLS regressions. In columns 1, 3 and 5, the dependent variable is the student's answer to a 5-point Likert confidence question, which asked whether they thought they had the academic aptitude and skills to graduate from college (column 1)), to graduate with an Industrial Engineering degree (column 3) and graduate with a Mechanical Engineering degree (column 5). For ease of interpretation, the answers have been standardized around the control mean, and the estimates are therefore expressed in standard deviations from such mean. In column 2, 4 and 6, we report estimated impact on a 0-1 indicator for the student having answers "definitely yes" to the corresponding question. The number of observations in columns 1 and 2 is lower than 490 because the confidence questions regarding completing college was only asked to students who stated that they intended to go to college after high school graduation. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

Table 8: Mechanisms: Impacts of Parents, Teachers' and Peers' Preferences and Support

	Mom's preferences		Dad's preferences		Teachers' support		Peers' preferences	
	College	Engineering	College	Engineering	College	Engineering	College	Engineering
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instagram (T1)	0.001	-0.025	0.034*	-0.053	-0.018	-0.026	0.067**	-0.085
	(0.018)	(0.025)	(0.018)	(0.045)	(0.046)	(0.045)	(0.031)	(0.055)
Instagram & WhatsApp (T2)	0.011	-0.031	0.030	-0.039	-0.055	-0.002	0.057*	-0.114**
	(0.013)	(0.028)	(0.020)	(0.033)	(0.038)	(0.047)	(0.033)	(0.044)
Observations	490	490	490	490	490	490	490	490
Control Mean	0.973	0.144	0.938	0.192	0.699	0.322	2.897	0.534
Randomization Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LASSO Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.540	0.789	0.797	0.681	0.382	0.610	0.633	0.496

Notes: The table displays ITT estimates generated by OLS regressions. In columns 1 and 2, the dependent variables are 0-1 dummy variables equal to 1 if the student answered "go to college" and "study engineering", respectively, when asked what they thought their mother would want them to after high school graduation. The "Dad's preferences" outcome variables, in columns 3 and 4, are generated by the same questions asked with reference to fathers. The "Teachers' support" outcome variables, in columns 5 and 6, are constructed in a slightly different way. We asked student whether theu had discussed going to college with any of their teachers. The outcome variable in column 5 is a 0-1 indicator equal to 1 if the student answered affirmatively. We then asked students what was the field of study that teachers recommended the most. The outcome variable in column 6 is a 0-1 indicator, which is equal to 1 if the student answered "engineering." To measure perceptions of peers' preferences for college and for engineering, we asked students to list and name their three best friends. We then asked if each friend would go to college after graduation, and, if so, what their field of study will be. The outcome variable in column 7 is the number of friends, out of 3, who the student thinks will go to college. The outcome of variable in column 8 is the number of friends, out of 3, whom the student thinks will study engineering. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

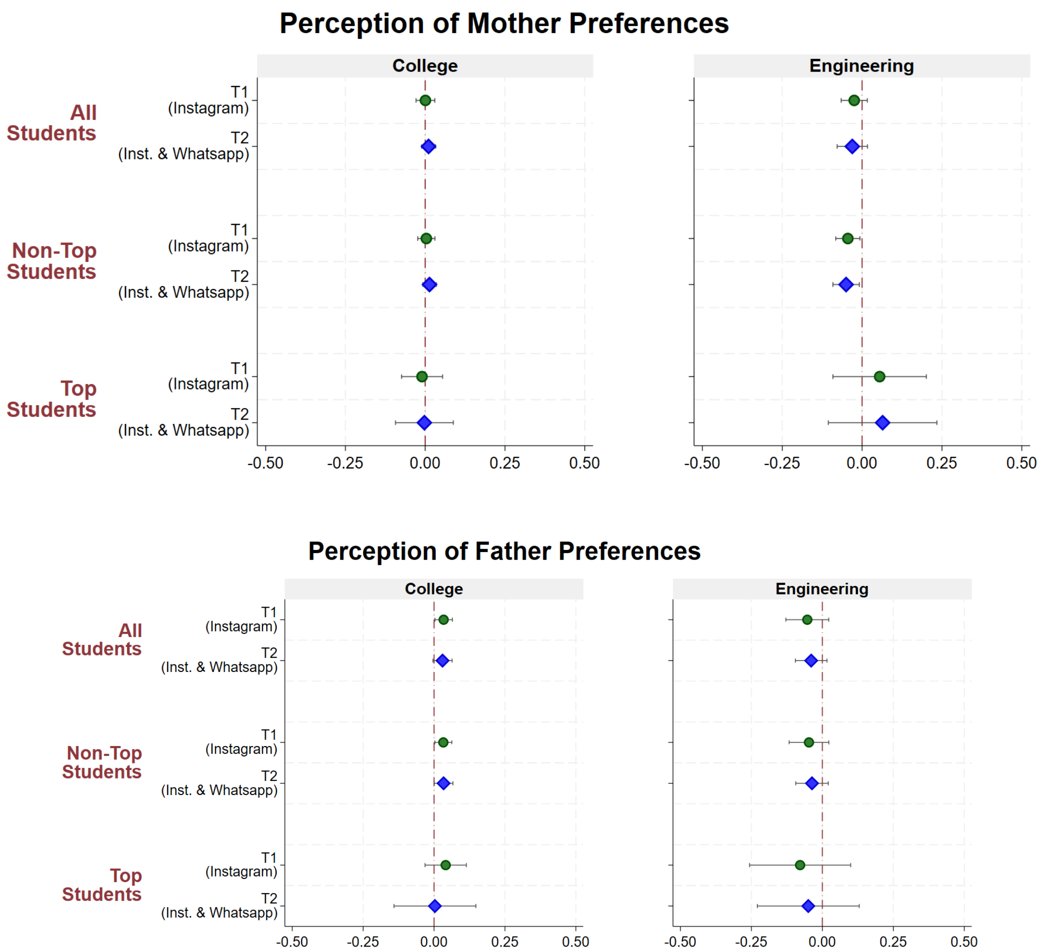
APPENDIX

Additional Figures and Tables

Figure A1: Map of the Regions of Peru

Notes: The map shows all of the regions of Peru. The red circle identifies the 6 regions of Northern Peru where the study took place.

Figure A2: Treatment Impacts on Perceptions of Parents' Preferences



Notes: Coefficient estimates come from ITT regressions reported in Tables 8 and A6. Controls are selected through the Double Lasso procedure from the set of variables displayed in Table 2, and always include a school-level indicators for low SES status, and co-educational status, in addition to metropolitan area fixed effects. The error bars correspond to 95% confidence intervals.

Table A1: Balance Tables - Panels by Gender

	Girls							Boys							
	Control	T1	T2	C=T1	C=T2	T1=T2	C=T1=T2	Control	T1	T2	C=T1	C=T2	T1=T2	C=T1=T2	(M=F)
Panel A: Demographics															
Age	16.08	16.22	16.12	0.006	0.292	0.026	0.008	16.18	16.12	16.17	0.358	0.876	0.479	0.677	0.536
	(0.46)	(0.45)	(0.35)					(0.39)	(0.33)	(0.45)					
Mom Went to College	0.30	0.24	0.34	0.267	0.433	0.048	0.141	0.38	0.36	0.27	0.834	0.144	0.244	0.292	0.304
	(0.46)	(0.43)	(0.47)					(0.49)	(0.48)	(0.45)					
Dad Went to College	0.33	0.24	0.33	0.065	0.955	0.046	0.089	0.36	0.31	0.29	0.527	0.323	0.792	0.603	0.522
	(0.47)	(0.43)	(0.47)					(0.48)	(0.47)	(0.46)					
Math Performance	16.56	16.73	16.34	0.437	0.328	0.052	0.172	16.88	17.08	16.87	0.605	0.961	0.554	0.805	0.014
	(2.21)	(1.96)	(1.97)					(1.92)	(2.55)	(1.95)					
Language Performance	17.33	17.30	17.23	0.847	0.570	0.703	0.848	16.75	16.39	16.96	0.292	0.414	0.067	0.165	0.000
	(1.78)	(1.70)	(1.51)					(1.73)	(2.34)	(1.60)					
Wish College	0.87	0.87	0.83	0.998	0.289	0.271	0.436	0.84	0.83	0.75	0.799	0.120	0.228	0.230	0.051
	(0.34)	(0.34)	(0.38)					(0.37)	(0.38)	(0.44)					
Wish Engineering	0.14	0.13	0.13	0.740	0.741	0.994	0.931	0.36	0.39	0.28	0.744	0.256	0.153	0.309	0.000
	(0.35)	(0.34)	(0.34)					(0.48)	(0.49)	(0.45)					
Wish STEM	0.17	0.15	0.14	0.589	0.392	0.753	0.687	0.39	0.41	0.32	0.842	0.363	0.283	0.497	0.000
	(0.38)	(0.36)	(0.35)					(0.49)	(0.50)	(0.47)					
Panel B: Self-Reported Confidence															
Being in Top 50%	0.32	0.29	0.29	0.487	0.595	0.855	0.770	0.45	0.58	0.49	0.146	0.597	0.302	0.337	0.000
	(0.47)	(0.45)	(0.46)					(0.50)	(0.50)	(0.50)					
Being in Top 20%	0.33	0.24	0.24	0.065	0.066	0.960	0.103	0.43	0.44	0.36	0.916	0.384	0.349	0.563	0.000
	(0.47)	(0.43)	(0.43)					(0.50)	(0.50)	(0.48)					
Completing College	4.38	4.35	4.38	0.756	0.947	0.694	0.918	4.47	4.53	4.58	0.529	0.237	0.628	0.483	0.003
	(0.73)	(0.72)	(0.74)					(0.65)	(0.56)	(0.57)					
Studying Industrial Engineering	3.06	3.00	2.85	0.681	0.115	0.260	0.275	3.65	3.50	3.54	0.428	0.498	0.847	0.697	0.000
	(1.19)	(1.31)	(1.21)					(1.07)	(1.15)	(1.13)					
Studying Mechanical Engineering	2.72	2.63	2.54	0.505	0.156	0.457	0.371	3.44	3.42	3.26	0.921	0.320	0.414	0.551	0.000
	(1.23)	(1.25)	(1.18)					(1.14)	(1.22)	(1.21)					
Panel C: Perceptions															
Mom Wishes College	0.98	0.94	0.98	0.053	0.919	0.046	0.045	0.96	0.95	0.96	0.818	0.962	0.843	0.970	0.542
	(0.14)	(0.24)	(0.14)					(0.19)	(0.21)	(0.20)					
Dad Wishes College	0.98	0.91	0.95	0.009	0.303	0.067	0.017	0.95	0.95	0.94	0.891	0.807	0.710	0.927	0.969
	(0.16)	(0.29)	(0.21)					(0.22)	(0.21)	(0.24)					
Mom Wishes Engineering	0.14	0.10	0.12	0.264	0.637	0.493	0.535	0.27	0.34	0.26	0.365	0.881	0.270	0.508	0.000
	(0.35)	(0.30)	(0.33)					(0.45)	(0.48)	(0.44)					
Dad Wishes Engineering	0.18	0.17	0.17	0.883	0.726	0.835	0.939	0.34	0.34	0.32	0.940	0.841	0.787	0.960	0.000
	(0.38)	(0.38)	(0.37)					(0.48)	(0.48)	(0.47)					
Teacher Wishes College	0.58	0.74	0.70	0.002	0.018	0.446	0.006	0.73	0.64	0.71	0.272	0.770	0.377	0.515	0.591
	(0.50)	(0.44)	(0.46)					(0.45)	(0.48)	(0.46)					
Teacher Wishes Engineering	0.08	0.12	0.10	0.234	0.624	0.448	0.474	0.22	0.27	0.26	0.539	0.524	0.966	0.774	0.000
	(0.27)	(0.32)	(0.29)					(0.42)	(0.45)	(0.44)					
Gender Attitudes	-0.00	0.09	0.08	0.260	0.270	0.912	0.446	-0.19	-0.33	-0.15	0.317	0.743	0.173	0.378	0.000
	(0.68)	(0.73)	(0.66)					(0.80)	(0.82)	(0.80)					
Self-Efficacy	-0.00	0.03	0.00	0.248	0.854	0.301	0.444	0.09	0.11	0.12	0.702	0.540	0.820	0.814	0.000
	(0.25)	(0.25)	(0.25)					(0.25)	(0.23)	(0.28)					
Panel D: Social Media															
Has IG Account	0.87	0.85	0.90	0.563	0.468	0.174	0.397	0.75	0.70	0.71	0.508	0.498	0.957	0.745	0.000
	(0.34)	(0.36)	(0.31)					(0.43)	(0.46)	(0.46)					
Spends >2hr on social media	0.47	0.46	0.49	0.857	0.694	0.550	0.829	0.32	0.34	0.32	0.813	0.984	0.787	0.959	0.000
	(0.50)	(0.50)	(0.50)					(0.47)	(0.48)	(0.47)					
Followed Account	0.00	0.27	0.32	0.000	0.000	0.287	0.000	0.00	0.02	0.03	0.274	0.125	0.557	0.299	0.000
	(0.00)	(0.45)	(0.47)					(0.00)	(0.12)	(0.17)					
Observations	162	185	200					77	64	99					

Notes: The table displays averages for respondents by treatment status and gender. Standard deviations are in parentheses. For each panel and each variable, the p-values disclosed in the last three columns are generated by pairwise tests of equality of means. The p-value disclosed in the last column of the table is generated by a test of equality of means between all boys and all girls in the sample.

Table A2: Attrition

	Attrited at Endline	
	(1)	(2)
Instagram (T1)	-0.031	-0.036
	(0.045)	(0.047)
Instagram & WhatsApp (T2)	0.014	0.021
	(0.043)	(0.041)
Low SES School		0.053
		(0.061)
Co-ed School		-0.044
		(0.043)
Baseline Interest in College		-0.028
		(0.029)
Baseline Interest in Engineering		-0.013
		(0.034)
Top Math Student		0.059
		(0.053)
Math Score		-0.010
		(0.011)
Language Score		0.007
		(0.011)
Observations	547	547
Clusters	73	73
Control Mean	0.901	0.901
Controls	No	Yes
$H_0 : T1 = T2$	0.318	0.184

Notes: OLS regressions. Robust standard errors clustered at the school level in parentheses. The sample has been restricted to girls. The outcome in both columns is a binary indicator equal to 1 if a student completed the follow-up survey, and is 0 if a student did not complete the follow-up survey. "Low SES School" is a binary indicator equal to 1 if a student attended a low socio-economic status school, and is 0 otherwise. "Co-ed School" is a binary indicator equal to 1 if a student attended a co-ed school, and is 0 otherwise. Baseline interest in engineering and college are both binary indicators generated from the baseline survey on interest in engineering and college, respectively. Math and Language Scores are a student's scores on math and language performance exams. Top Math Score is a binary indicator for having scored in the top 20% of the math performance exam.

Table A3: Treatment Impacts on Boys

	Wish Engineering	Scholarship	Enrolled in Engineering	Enrolled in College
	(1)	(2)	(3)	(4)
Instagram (T1)	0.029	-0.064	0.023	0.095
	(0.044)	(0.061)	(0.050)	(0.090)
Instagram & WhatsApp (T2)	-0.034	0.050	-0.000	-0.001
	(0.033)	(0.057)	(0.047)	(0.099)
Observations	212	240	240	240
Clusters	49	55	55	55
Control Mean	0.352	0.156	0.182	0.416
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	–	–	–
LASSO Controls	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.211	0.061	0.698	0.363

Notes: The table displays ITT estimates from linear probability models, where the dependent variables are our four primary outcomes of interest. The sample is restricted to boys. A total of 240 boys participated in the baseline survey. Of them, 212 participated in the endline survey. We have scholarship application and enrollment data for the full sample. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. When available, we include the outcome variable measured at baseline. Other controls in are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table A1. Standard errors clustered at the school level are reported in parentheses.

Table A4: Confidence in Pursuing Degrees Other Than Engineering

	Law		Medicine		Business		Economics		Education	
	Confidence	Def. yes	Confidence	Def. yes	Confidence	Def. yes	Confidence	Def. yes	Confidence	Def. yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Instagram (T1)	0.131	0.031	-0.033	0.042	-0.094	-0.019	-0.076	0.015	-0.039	0.027
	(0.100)	(0.050)	(0.079)	(0.043)	(0.114)	(0.040)	(0.099)	(0.035)	(0.089)	(0.056)
Instagram & WhatsApp (T2)	-0.022	-0.030	0.001	0.074*	0.022	0.033	0.082	0.058**	-0.099	0.044
	(0.106)	(0.041)	(0.077)	(0.044)	(0.110)	(0.042)	(0.086)	(0.028)	(0.081)	(0.042)
Observations	490	490	490	490	490	490	490	490	490	490
Control Mean	-0.000	0.308	0.000	0.212	0.000	0.308	-0.000	0.171	0.000	0.322
Randomization Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LASSO Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.107	0.207	0.677	0.390	0.313	0.139	0.067	0.166	0.447	0.688

Notes: The table displays ITT estimates generated by OLS regressions. In columns 1, 3, 5, 7 and 9, the dependent variable is the student's answer to a 5-point Likert confidence question, which asked whether they thought they had the academic aptitude and skills to graduate to graduate from college with a degree in the field indicated at the top of the corresponding column. For ease of interpretation, the answers have been standardized around the control mean, and the estimates are therefore expressed in standard deviations from such mean. In column 2, 4 and 6, 8 and 10, we report estimated impacts on a 0-1 indicator for the student having answers "definitely yes" to the corresponding question. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

Table A5: Self-Reported Confidence. Top Student Heterogeneity

	Completing College		Industrial Engineering		Mechanical Engineering	
	Confidence	Def. yes	Confidence	Def. yes	Confidence	Def. yes
	(1)	(2)	(3)	(4)	(5)	(6)
Instagram (T1)	0.121	0.083	-0.067	-0.035	-0.118	-0.046
	(0.122)	(0.054)	(0.080)	(0.044)	(0.097)	(0.041)
Instagram & WhatsApp (T2)	0.065	-0.012	-0.070	0.007	-0.084	0.002
	(0.104)	(0.057)	(0.094)	(0.039)	(0.095)	(0.036)
Top	-0.154	-0.214*	-0.066	0.050	0.032	-0.004
	(0.255)	(0.111)	(0.252)	(0.115)	(0.181)	(0.070)
T1 × Top	0.128	0.132	0.253	0.053	0.140	0.025
	(0.236)	(0.117)	(0.241)	(0.117)	(0.158)	(0.089)
T2 × Top	-0.014	0.153	0.272	-0.074	0.149	-0.065
	(0.261)	(0.126)	(0.231)	(0.122)	(0.196)	(0.090)
Observations	463	463	490	490	490	490
Control Mean	-0.004	0.595	-0.083	0.145	-0.090	0.137
Randomization Controls	Yes	Yes	Yes	Yes	Yes	Yes
LASSO Controls	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.602	0.063	0.972	0.147	0.697	0.075
$H_0 : T1 + T1 \times Top = 0$	0.143	0.036	0.357	0.860	0.866	0.787
$H_0 : T2 + T2 \times Top = 0$	0.816	0.202	0.285	0.542	0.721	0.456
$H_0 : T1 + T1 \times Top = T2 + T2 \times Top$	0.280	0.490	0.926	0.510	0.761	0.637

Notes: The table displays ITT estimates generated by OLS regressions. In columns 1, 3 and 5, the dependent variable is the student's answer to a 5-point Likert confidence question, which asked whether they thought they had the academic aptitude and skills to graduate from college (column 1)), to graduate with an Industrial Engineering degree (column 3) and graduate with a Mechanical Engineering degree (column 5). For ease of interpretation, the answers have been standardized around the control mean, and the estimates are therefore expressed in standard deviations from such mean. In column 2, 4 and 6, we report estimated impact on a 0-1 indicator for the student having answers "definitely yes" to the corresponding question. The number of observations in columns 1 and 2 is lower than 490 because the confidence questions regarding completing college was only asked to students who stated that they intended to go to college after high school graduation. "Top" indicates whether the student was among the top performers (top quantile) in math, based on their final grade on the year preceding the study, i.e, the year before high school graduation. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

Table A6: Perceptions. Top Student Heterogeneity

	Mom's preferences		Dad's preferences		Teachers' support		Peers' preferences	
	College	Engineering	College	Engineering	College	Engineering	College	Engineering
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instagram (T1)	0.003	-0.044*	0.029	-0.043	-0.014	-0.040	0.080*	-0.084
	(0.016)	(0.024)	(0.018)	(0.040)	(0.053)	(0.052)	(0.041)	(0.068)
Instagram & WhatsApp (T2)	0.015	-0.050*	0.032	-0.029	-0.047	-0.018	0.052*	-0.116**
	(0.012)	(0.025)	(0.020)	(0.032)	(0.042)	(0.055)	(0.031)	(0.053)
Top	-0.008	-0.022	-0.053	0.057	0.105	0.102	0.006	-0.096
	(0.024)	(0.069)	(0.048)	(0.084)	(0.084)	(0.132)	(0.045)	(0.084)
T1 × Top	-0.013	0.101	0.017	0.008	0.004	0.113	-0.077	0.009
	(0.034)	(0.094)	(0.044)	(0.099)	(0.086)	(0.137)	(0.091)	(0.119)
T2 × Top	-0.022	0.102	-0.032	0.007	0.002	0.092	-0.019	0.054
	(0.056)	(0.100)	(0.088)	(0.111)	(0.107)	(0.174)	(0.050)	(0.167)
Observations	490	490	490	490	490	490	490	490
Control Mean	0.974	0.145	0.940	0.171	0.667	0.299	2.889	0.547
Randomization Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LASSO Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.438	0.775	0.879	0.654	0.458	0.651	0.357	0.545
$H_0 : T1 + T1 \times Top = 0$	0.813	0.528	0.316	0.744	0.886	0.544	0.969	0.428
$H_0 : T2 + T2 \times Top = 0$	0.887	0.614	0.999	0.835	0.667	0.636	0.463	0.673
$H_0 : T1 + T1 \times Top = T2 + T2 \times Top$	0.980	0.946	0.560	0.906	0.697	0.995	0.594	0.937

Notes: The table displays ITT estimates generated by OLS regressions. In columns 1 and 2, the dependent variables are 0-1 dummy variables equal to 1 if the student answered "go to college" and "study engineering", respectively, when asked what they thought their mother would want them to after high school graduation. The "Dad's preferences" outcome variables, in columns 3 and 4, are generated by the same questions asked with reference to fathers. The "Teachers' support" outcome variables, in columns 5 and 6, are constructed in a slightly different way. We asked student whether theu had discussed going to college with any of their teachers. The outcome variable in column 5 is a 0-1 indicator equal to 1 if the student answered affirmatively. We then asked students what was the field of study that teachers recommended the most. The outcome variable in column 6 is a 0-1 indicator, which is equal to 1 if the student answered "engineering." To measure perceptions of peers' preferences for college and for engineering, we asked students to list and name their three best friends. We then asked if each friend would go to college after graduation, and, if so, what their field of study will be. The outcome variable in column 7 is the number of friends, out of 3, who the student thinks will go to college. The outcome of variable in column 8 is the number of friends, out of 3, whom the student thinks will study engineering. "Top" indicates whether the student was among the top performers (top quantile) in math, based on their final grade on the year preceding the study, i.e, the year before high school graduation. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

Table A7: Heterogeneity by School Co-ed Status

	Wish Engineering	Scholarship	Enrolled in Engineering	Enrolled in College
	(1)	(2)	(3)	(4)
Instagram (T1)	0.030	-0.142***	-0.001	0.199
	(0.026)	(0.035)	(0.023)	(0.132)
Instagram & WhatsApp (T2)	-0.041**	-0.051*	0.037	0.124**
	(0.020)	(0.028)	(0.033)	(0.051)
Co-ed School	0.053*	-0.185***	-0.018	-0.017
	(0.031)	(0.032)	(0.027)	(0.071)
T1 × Co-ed School	-0.162	0.220	0.020	0.004
	(0.040)	(0.048)	(0.039)	(0.159)
T2 × Co-ed School	-0.070	0.148	-0.039	-0.024
	(0.039)	(0.052)	(0.037)	(0.087)
Observations	490	547	547	547
Clusters	67	73	73	73
Control Mean	0.143	0.200	0.077	0.292
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	–	–	–
LASSO Controls	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.010	0.006	0.299	0.548
$H_0 : T1 + T1 \times$ Co-ed $= 0$	0.000	0.007	0.490	0.016
$H_0 : T2 + T2 \times$ Co-ed $= 0$	0.001	0.017	0.928	0.153
$H_0 : T1 + T1 \times$ Co-ed $= T2 + T2 \times$ Co-ed	0.282	0.643	0.389	0.162

Notes: The table displays ITT estimates from linear probability models, where the dependent variables are our four primary outcomes of interest. "Co-ed School" is a 0-1 dummy, equal to 1 if the school is mixed-gender and 0 if the school only serves girls. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.

Table A8: Heterogeneity by School Socio-Economic Status (SES) Classification

	Wish Engineering	Scholarship	Enrolled in Engineering	Enrolled in College
	(1)	(2)	(3)	(4)
Instagram (T1)	-0.078*	-0.010	0.009	0.195**
	(0.042)	(0.045)	(0.019)	(0.097)
Instagram & WhatsApp (T2)	-0.085**	0.070	0.032*	0.156**
	(0.036)	(0.048)	(0.018)	(0.068)
High-SES School	-0.023	-0.000	0.112***	0.373***
	(0.041)	(0.043)	(0.031)	(0.093)
T1 $\times$ High-SES School	0.033	0.019	0.021	0.067
	(0.057)	(0.069)	(0.054)	(0.132)
T2 $\times$ High-SES School	0.029	-0.084	-0.077	-0.142
	(0.051)	(0.060)	(0.038)	(0.112)
Observations	490	547	547	547
Clusters	67	73	73	73
Control Mean	0.221	0.117	0.039	0.126
Randomization Controls	Yes	Yes	Yes	Yes
Baseline Value Controls	Yes	–	–	–
LASSO Controls	Yes	Yes	Yes	Yes
$H_0 : T1 = T2$	0.819	0.094	0.307	0.626
$H_0 : T1 + T1 \times$ High-SES $= 0$	0.108	0.834	0.524	0.001
$H_0 : T2 + T2 \times$ High-SES $= 0$	0.071	0.629	0.172	0.867
$H_0 : T1 + T1 \times$ High-SES $= T2 + T2 \times$ High-SES	0.706	0.521	0.129	0.001

Notes: The table displays ITT estimates from linear probability models, where the dependent variables are our four primary outcomes of interest. "High SES School" is a 0-1 dummy, equal to 1 if the school is classified as High Socio-Economic Status, and 0 otherwise. In all regressions, we control for the randomization strata, which are municipalities, whether the school is co-ed and whether the school is classified as a low SES school. We also include the outcome variable measured at baseline. Additional controls are selected using the Double Lasso procedure (Belloni et al., 2014) from the full set of covariates displayed in Table 2. Standard errors clustered at the school level are reported in parentheses.