To inspire and to inform: The role of role models*

Marcos Agurto[†], Muchin Bazan[‡], Siddharth Hari[§], Sudipta Sarangi[¶]

Abstract

Bridging the gender gap in STEM fields has emerged as a concern for policymakers both in developed and developing countries. This paper examines the impact of light touch intervention where female engineering students act as role-models for high school students in Peru. We find that a brief 20-minute interaction with the role-models led to sharp increases in preferences towards engineering, with the effects being concentrated on female students with high math aptitude. We find that these results are driven by increased self-confidence as a result of exposure to role models. Set in the context of a developing country, our results show that low-cost interventions can be helpful in reducing the STEM gender gap but cannot address broader deep seated gender stereotypes.

Keywords: STEM gender gap, role models, career choices, stereotypes, RCTs.

JEL Classification codes: C93, I23, I24, J16.

---

*We would like to thank seminar participants at Virginia Tech, 8th UDEP Workshop for Young Economists, SEA 90th Annual Meeting, 47th Annual Conference Eastern Economic Association, and 2021 Global Labor Organization Young Scholar Program Meeting. We also thank Jorge Agüero, Niloy Bose, Suqin Ge, Xu Lin, Subha Mani, Melinda Miller, and Kompal Sinha for their suggestions to improve this paper. Sarangi would like to acknowledge the support of Collegium de Lyon where a part of this research was carried out. The unique identifying number for the AEA registry is: AEARCTR-0002945. The opinions expressed in the paper are those of the authors and do not necessarily reflect the views of the World Bank Group or its Board of Executive Directors.

[†]The Department of Economics. Universidad de Piura and Lima School of Economics. Av. Ramon Mugica 131, Urb. San Eduardo, Piura, Peru. Email: marcos.agurto@udep.edu.pe.

[‡]The Department of Economics, Virginia Tech and Universidad de Piura, GLO Affiliate. 3016 Pamplin Hall, 880 West Campus Drive, Blacksburg, VA 24061, United States. Email: muchin90@vt.edu.

[§]The World Bank. 1818 H Street, NW Washington, DC 20433 USA. Email: shari@worldbank.org.

[¶]Corresponding author. The Department of Economics, Virginia Tech and Collegium de Lyon. 3016 Pamplin Hall, 880 West Campus Drive, Blacksburg, VA 24061, United States. Email: ssarangi@vt.edu.

1 Introduction

Despite significant progress in women's access to college education in recent decades, there are still large gender enrollment gaps in Science, Technology, Engineering, and Mathematics (STEM) majors. At the tertiary level women are 13 percentage points less likely to enroll in STEM fields in low-middle income countries, a gap which rises to 17 percentage points in high income countries. Further, within STEM disciplines, women's participation in engineering, manufacturing and construction related fields globally stands at 7 percent for women compared to 22 percent for men [World Bank (2020)]. The gender gap in engineering has profound consequences for women in particular and society in general. It contributes to the under-representation of females at the top of the income distribution,[1] has adverse effects on the development of new ideas, science, technology and firm productivity, and negative repercussions on economic growth via the misallocation of talent.[2]

This paper addresses an important, yet under-studied factor that can play a key role in explaining the gender gap in STEM careers: the lack of appropriate female role models. Role models can play an important role in shaping career choices of adolescents by providing relevant information and being sources of inspiration. One may imagine that due to social conditioning female students may believe that STEM careers are not for them. However, when presented with role models in these professions that may revise the prior beliefs about being able to succeed in these professions.

However, several barriers restrict young girls from having female role models in STEM, including their relative scarcity in many contexts, as well as the absence of initiatives that bring the experience of the few existing ones to young girls. It is therefore difficult for female students to come into direct contact with women who have majored in or are majoring in STEM fields. Working with a private university in Peru, Universidad de Piura (UDEP), we designed a light-touch intervention (Rury & Carrell (2023)) that aims to bring those relatively scarce role models close to high school girls and examine the impact of such an intervention on college major

[1] See for instance, Blau & Kahn (2017).

[2] Detailed discussions can be found in Bear & Woolley (2011), and Hsieh et al. (2019).

choices.[3] To the best of our knowledge, this is the first paper studying the impact of role-models in addressing the STEM gender gap in the context of a developing country.

Our field experiment involved classroom talks at randomly selected schools to senior-year high school students, who were about to make decisions related to college enrollment and field of study. These talks were given by female students currently enrolled in an engineering major (or recent graduates), with first-hand experience and knowledge on the skills, aptitude and motivation needed to successfully pursue a degree in engineering.

We exploit this experimental setting to test whether exposure to female senior students/recent graduates in engineering acting as role models can change high school girls' perceptions regarding STEM fields, self-confidence and ultimately influence choice of college majors. The task of the role models was to deliver a 20-minute motivational presentation in treatment schools and to answer questions thereafter. Overall, the main message transmitted to students can be summarized as follows:

> *"You do not need to be mathematical genius to become a successful engineer"*, *"Boys and girls have the same intellectual capacity even though their brains are physically different"*, *"women are very creative and they can contribute to new ideas"*,*"To study Engineering creativity, ingenuity, effort, and desire to change the world are also very important"*.

Our sample consists of 5,378 students in the final year of high school across 51 treated and 58 control group schools. We find that our treatment has strong effects for female students with high quantitative ability. In particular, while the aggregate treatment effects are small and statistically indistinguishable from zero, for female students in the top quartile of the (baseline) math score distribution, exposure to role models increases the likelihood of planning to enroll in engineering by 9 percentage points. Given the low baseline levels of preferences towards engineering, this translates to a 45 percent increase. We find no effects for other female students. We also find weak evidence suggesting that boys in the two lowest math ability

[3] UDEP is a private university in Peru. It ranks among the top ten universities in Peru according to the QS Latin American University Rankings 2023.

quartiles increase their preferences for engineering majors (although these are not statistically significant).

We then explore the possible mechanisms that could be driving these results. We find that for girls with high quantitative ability, exposure to role models significantly increases self-confidence in their aptitude and skills to do well in engineering majors. However, we find no effect of the intervention on gender stereotypes related to STEM fields overall. When presented with information about a hypothetical successful engineer and asked to guess their gender, we find that our treatment has no effect on boys or girls. Thus, not surprisingly, when asked what major they would recommend to a hypothetical female friend ("Lorena") and a hypothetical male friend ("Javier"), once again we find no effects of our exposure to role models. These results indicate that broader gender based stereotypes may be deep rooted which are not likely to be affected by a 20 minute interaction, and would require a stronger and more likely a prolonged intervention

In the case of boys, the mechanisms behind the (weaker) results are less clear; but the evidence suggests that they are related to the specific information provided in the talks, which stressed that skills other than math ability are also relevant in engineering majors. It is also important to highlight that our results regarding the potential mechanisms are stronger among students residing within the UDEP geographical area of influence (i.e in cities closer to UDEP where students have greater familiarity with the university). Quite likely high math ability female students that live further away from UDEP and are not familiar with how the university functions may not be as easily persuaded by the message of the role models. The importance of geographic proximity mentoring interventions is in-line with in previous studies in the literature. [4] Taken together, all these suggest that context is important for creating effective role models. These results also suggest the potential role for *light touch* interventions to increase participation of women in STEM fields . Such interventions have recently been shown to have strong effects on educational outcomes

[4] For example Hardt et al. (2020) studies the effect of online mentoring programs at a public German university to improve online teaching effectiveness during COVID-19. The mentees in this study were undergraduate students enrolled in the second term, and mentors were more senior students but enrolled in the same study program as the mentees. The authors find that students who reside in the region where the university is located benefit more from the program. The mentoring program had positive effects on students' motivation, exam registrations, and academic performance.

of underrepresented groups (Rury & Carrell (2023)).

Our paper contributes to three strands of the literature. First, it adds to the extensive body of research on the causes of the STEM gender gap, particularly in engineering. Understanding the causes of low female participation in STEM fields continues to be an important research and policy question that has been studied by professionals in various fields. Several factors have been analyzed as potential determinants of the observed gender STEM gap. These include examining the role of differences in biological characteristics [Ellison & Swanson (2010), Nollenberger et al. (2016), UNESCO (2017)] as well as non-cognitive abilities (i.e. self-efficacy, self-perception) [Correll (2001), Eble & Hu (2020), Kahn & Ginther (2017)]. Within STEM fields, the evidence strongly suggests that engineering has a low proportion of females mainly due to women's perception that engineering is a career not suitable for them [Emerson et al. (2012)]. Moreover, these stereotypes are transmitted to girls from a young age. Girls consider engineering as a "masculine" domain and believe that women cannot succeed there [Reuben et al. (2017), Eble & Hu (2022)], as they lack the necessary skills as well as due to the discrimination faced by women from STEM fields in the labor market [Bayer & Rouse (2016)]. Additionally, some authors point out the competitive nature of STEM fields as a reason for the low participation of women in Science and Engineering careers [Gneezy et al. (2003), Niederle & Vesterlund (2007), Niederle & Vesterlund (2010), Buser et al. (2014), Flory et al. (2014)]. Our study shows that brief interactions with external (non-teaching) female students or recent graduates in engineering disciplines can influence perceptions, self-confidence and ultimately career choices of a certain segment of female students.

Second, we contribute to the literature that assesses the effectiveness of mentors and role-models on students' academic performance, enrollment and drop-out decisions, and occupational choices. This literature primarily focuses on the role of teachers or instructors.[5] However, many of these studies suffer from identification issues related to the unobserved

[5] Although this is not an exhaustive list, a wide range of issues relevant to this aspect can be found in the papers by Neumark & Gardecki (1998), Bettinger & Long (2005), Dee (2007), Hoffmann & Philip (2009), Carrell et al. (2010), Bottia et al. (2015), Eble & Hu (2020), Lim & Meer (2017), Kofoed & McGovney (2019), and Lim & Meer (2019).

preferences of instructors towards same gender students [Zeltzer (2020)], as well as the self-selection of students choosing to attend classes with instructors who they like the most, or have less strict grading policies. As a result of these constraints, causality cannot be clearly established. Other studies have also analyzed the effects of non-teaching role model interventions in the field [Beaman et al. (2012), Del Carpio & Guadalupe (2018), Ashraf et al. (2020), Porter & Serra (2020), Breda et al. (2020), Brooks et al. (2018), Lafortune et al. (2018)]. Most closely related to our work are two papers. The first is a recent study by Porter & Serra (2020), which studies the effects of exposure of freshman undergraduate students in the US to professional economists acting as role models and finds an increase in enrollment in the economics major. The second is a paper by Breda et al. (2020) where middle-aged female role models (scientists and PhD students) were able to influence French high school students' perception towards STEM fields.

Our work differs in important ways. First, to the best of our knowledge, this is the first paper which examines the impact of role-models on addressing the gender gap in STEM fields in the context of a developing country. Promoting participating in STEM fields is particularly important in developing countries where the gender gap is wider so that they can take advantage of the jobs of the future, driven by technological advances [World Bank (2020)], to minimize the misallocation of talent and to reduce pay inequities. Second, the role models in our intervention are senior college students or very recent graduates in engineering, and therefore younger and closer in age to the target group than the role models in Breda et al. (2020) and Porter & Serra (2020). Therefore, we expect our role models not only to motivate high school students, but also for high school students to feel more connected to them.

Third, the paper contributes to studies in social psychology that look at the effect of gender stereotypes on women's under-representation in science. Several studies in social psychology have analyzed mentoring programs and non-teaching role model interventions (see for instance Macphee et al. (2013)), but have not been successful at tracking the causal effects on career choices and isolating the related mechanisms. Our paper fills this gap as well.

The remainder of the paper is organized as follows. Section 2 describes the Peruvian educational system, the status of females participation in STEM fields and the context and setting of our experiment. Section 3 presents the data and empirical strategy. Section 4 presents the intervention results and discusses potential mechanisms. Robustness checks are presented in section 5, while additional heterogeneous effects are discussed in section 6. Section 7 concludes and discusses policy implications of these findings.

2 The Experimental Setting

2.1 Peruvian Education System

The school system in Peru consists of six years of elementary education followed by five years of secondary education. School attendance in the country is compulsory from ages 5 to 16. Approximately 2.5 million students are enrolled at the high school level,[6] and 15,000 high schools are active across the 25 country regions.[7] At the high school level classes are usually administered by different instructors depending on the subject, and the school year runs from March to December.[8] While the government runs a public school system, for-profit and not-for-profit private schools also exist. The curriculum, which is defined by the Ministry of Education and must be followed by all schools in the country, does not distinguish between students who aim to pursue STEM and non-STEM college majors at any level of basic or high school education. Furthermore, Peru does not have a centralized university admission system and each university is responsible for its own admission process. In public universities, admission basically depends on a general examination test set by each university. In

[6] 76% of students are enrolled in a public school, and 90% of students are registered in schools located in urban areas. Source: 2017 Census of Schools, Ministry of Education (MINEDU), http://escale.minedu.gob.pe/resultado_censos.

[7] The Peruvian territory is divided into three administrative units: i) 25 regions, ii) 196 provinces, and 1,874 districts (municipalities). There are in total 8 provinces and 65 districts within the Piura region.

[8] Subjects that form part of the common National Curriculum are Mathematics, Communication, Foreign Language, Art, History, Geography, Economics, Civic, Social Skills, Physical Education, Religious Education, Science, Technology, and Environmental Studies.

some private universities, other admission mechanisms are also present.[9]

The schools in our intervention sample are located in 6 out of the 25 regions in Peru, all of them in the northern part of the country, as it can be observed in Figure 1. Roughly 60% of the schools (64 schools) are in the Piura Region, where UDEP's main campus is located. Of the remaining schools in our sample, 11 schools are located in La Libertad, 12 in Cajamarca, 3 in Ancash, 12 in Lambayeque and 7 in Tumbes.

While STEM careers cover various disciplines, in Peru engineering is by far the preferred STEM program among high school graduates. During the period 2016-2017, 93% of the roughly 417,000 students who applied for admission into a STEM field did so in engineering.[10] As in other countries around the world and in the Latin American region, in Peru, females are underrepresented in STEM fields in general and in engineering majors in particular. In this Andean country, during the period 2016-2017 only 30% (1 out 3) of those applying for admission into a STEM field[11] were women.[12] Moreover, while roughly just one in five (19%) female college applicants across the country selected engineering majors during this period; close to one in two (46%) male applicants chose an engineering program.

2.2 Universidad de Piura

UDEP is a not-for-profit private university located on the city of Piura, in the northern coast of Peru. According to recent national rankings, UDEP is one of the top 10 private universities in Peru, and the top ranked university in the northern region of the country. Historically, UDEP students come predominantly from the Piura region; however UDEP has also consistently attracted students from the neighbouring regions of Lambayeque (to the south) and Tumbes (to the north). Students from these three regions constitute about 95% of the UDEP Piura campus student population. In this sense, UDEP's prestige and reputation as a regional university

[9]For example, some private colleges offer direct admission to students in the upper third of their class GPA distribution.

[10]Administrative records of the Peruvian National Superintendence of Higher Education (SUNEDU): https://www.sunedu.gob.pe/sibe/.

[11]STEM fields include Biology, Mathematics, Statistics, Engineering, Physics, and Chemistry. Medical undergraduate studies, such as nursing and medicine, are not considered STEM in the Peruvian national statistics. In medical undergraduate studies, women are over-represented (70% are women).

[12]Administrative records of the Peruvian National Superintendence of Higher Education (SUNEDU): https://www.sunedu.gob.pe/sibe/.

is mainly concentrated in Piura region and the neighbouring regions of Tumbes and Lambayeque, which we refer to as UDEP's catchment area (and make up close to 80 percent of the schools in our sample).

Established in 1969, the UDEP Piura campus has approximately 6,500 undergraduate students across 15 academic programs. Within the category of STEM majors, the Piura campus only offers programs in Engineering.[13] In general, high school students application patterns at UDEP resemble those observed at the country level. According to the university administrative records, only 20% of all female applicants at UDEP selected engineering majors during the period 2016-2017. Moreover, 65% of engineering applicants were male, while just 35% were female.

2.3 The Field Experiment

Experimental design and randomization. The experiment started in early 2018 and was carried out in 18 cities.[14] These cities have a total of 225,000 high school students spread across 880 schools with women making up nearly half the student population.[15] Our team had access to a list of 150 schools within this area which have been frequently visited by UDEP admission officials in the last five years to promote the university and encourage applications. We finally chose 109 schools that make up our experimental sample,[16] which overall includes 5,378 students in the 11th grade.[17]

The randomization was stratified at the city level. Half of the schools in each city were assigned to the treatment group with the other half serving as controls. In total 51 and 58 schools were randomly assigned to treatment

[13] UDEP offers Engineering fields such as Civil Engineering, Industrial and System Engineering, and Mechanical and Electrical Engineering.

[14] The cities were Cajamarca (Cajamarca), Catacaos (Piura), Chiclayo (Lambayeque), Chimbote (Áncash), Chota (Cajamarca), Chulucanas (Piura), Cutervo (Cajamarca), La Union (Piura), Pacasmayo (La Libertad), Paita (Piura), Piura (Piura), Sechura (Piura), Sullana (Piura), Talara (Piura), Tambogrande (Piura), Trujillo (La Libertad), Tumbes (Tumbes), and Zarumilla (Tumbes).

[15] According to the 2019 Peruvian Ministry of Education School Census, this represents 33% of the high school student population in the Piura, Cajamarca, La Libertad, Lambayaque, Ancash and Tumbes regions, and 9% of the total high school enrollment in Peru. http://escale.minedu.gob.pe/padron-de-iiee.

[16] We excluded boys single-sex schools as well as schools outside Piura Region that could not be reached in a single bus trip.

[17] Power calculations where performed by the research team prior to the intervention. Based on a sample size of 5,450 students (109 clusters and 50 subjects per cluster), assuming an intra-cluster correlation of 0.05 with power of 80% we are able to detect an MDE of 0.19 standard deviations (with respect to the control group).

and control, respectively. Table 1 shows the baseline characteristics of students and schools by experimental group. As we can observe, our randomization was successful at achieving balance across control and treatment units observable characteristics.

The intervention. Role models visits took place between May and July 2018 and only targeted senior-year high school students. Our role models major in either i) civil engineering, or ii) industrial and systems engineering, or iii) mechanical and electrical engineering. They were 20 to 24 years old, and they were either engineering students in their fourth/fifth year of undergraduate studies or very recent graduates. In most cases, each treated school was visited by a single role model.[18]

It is worth mentioning that the role models prepared the presentation materials by themselves during several team-work sessions. They agreed on a general template, but adjustments were made to capture each role model's own experience as an engineering major. Role models also participated in a feedback session with UDEP faculty members before giving their talks. Most of the role models had previous experience in social events, group projects, and as volunteers in non-profit organizations and had developed strong communication skills.

Role models directly coordinated with UDEP Admissions Office on the date and time of the visits. The Admissions Office provided them with the school visits calendar and role models indicated the name of the person in charge of each talk. A lottery was used by the role models to assign the school visits conditional on each of them delivering approximately the same number of talks within and outside the Piura region. However, adjustments had to be made depending on role models' availability. On average, each role model delivered 5 talks.[19] Role models also received a monetary compensation, which was solely a function of the visited school's distance from Piura, and completely unrelated to performance in any sense. On average each role model received US$ 230 for their participation in the intervention.

[18] Two role models visited the treatment schools located in Trujillo and Tumbes. In these cases, only one role model did the presentation and the other accompanied her to the school visit. In seven other treatment schools, more than one role model gave the speech, which makes it difficult to identify the unique effect of each role model on the treated students. This happened because in these schools the number of sections was large and there were many senior-year students.

[19] Each role model gave between 3 and 7 talks.

It is important to highlight that in control schools, business continued as usual. That is, as in the last five years, these schools were visited by an UDEP admissions official who promoted all UDEP majors and admission mechanisms among senior high school students, without any mention of the role models intervention.

Content of the role model talks. During their presentations, which lasted approximately 20 minutes, role models used a set of slides highlighting the following facts: (i) gender differences in brain structure playing no role in determining males' and females' aptitude to pursue engineering majors, (ii) examples of contributions made by female engineers, (iii) definition of engineering as the art of solving problems and as a channel to change the world and make it better, (iv) statements aimed at deconstructing stereotypical views about engineering under the title "Beliefs or Reality?", (v) the experience of the role models at UDEP, and (vi) the relevance of creativity and ingenuity in engineering majors and the capability of girls to become engineers.[20] During and after their presentations, role models answered questions from students.

3 Data and Empirical strategy

3.1 Data

Student follow-up survey. In November 2018, four to six months after the role models' talks were delivered, we conducted a follow-up survey in 101 out of the 109 schools included in the experimental sample.[21] Students' responses to the survey were anonymous. [22][23]

The survey first asked students for their GPA scores in math, language

[20]In the talks the following statements were discussed: A person who wants to study engineering should be the top student in the class and a genius in mathematics, engineering is only for men, the engineers are boring, and women in engineering do not find jobs. Thumbnails of the slides shown during the school intervention are displayed in the Appendix. The full role models' presentation can be accessible throughout this link: https://drive.google.com/file/d/16VDemjA8wt2wY0-WGDMBHSMz-FycgLBp/view?usp=sharing.

[21]This included 54 out of 58 schools from the control group and 47 out of 51 schools from the treatment group. In 8 schools we could not conduct the follow-up survey because schools' authorities did not give us the necessary permission.

[22]Since we did not have access to the students' names or IDs, we are unable to match the survey data with UDEP's admissions or the Ministry of Education's administrative records.

[23]The survey was administered during class time; therefore, we have data on the students that were physically present at school on the day of the survey.

and science in the academic year that preceded our intervention; that is, when they were in 10th grade. We then asked students about the college major they would like to pursue, as well as several questions intended to measure self-confidence, gender beliefs, biases and perceptions. Regarding self-confidence, we asked students if they felt they have the abilities and skills needed to major in engineering at college. With respect to beliefs, biases and perceptions, we first asked students to imagine that they have two friends: "Javier", a boy; and "Lorena", a girl; and that both of them have a school GPA of 20 (the maximum possible score) in math and science. We then asked them which college major they would recommend to "Javier" and to "Lorena".[24] Similarly, we introduced to the students a hypothetical successful individual currently working in the engineering sector, and asked them whether that person was more likely to be a woman or a man. We also asked students to list at most five different engineering fields and collected information on students' expectations of the average monthly salary of a recent college graduate in engineering.

Finally, we collected data on parental demographic characteristics (i.e. age, education, working status, engineering background), siblings characteristics (i.e. number, gender, college major) and economic status (i.e. housing and other fixed assets ownership). Due to time, logistical and budget constraints, we did not survey students before the intervention. Therefore, we use the follow-up survey to capture information on students' pre-treatment characteristics. For such purpose, we will focus on socioeconomic variables which are unlikely to have changed over a 6 months period or have been affected by our intervention.

3.1.1 Data Analysis

Balance in observable individual and school level characteristics: We collected information on 5,378 senior-year high school students; 56% (2,998) of them are female and 50% (2,704) are in the treatment group. In Table 1 Panel A, we present the balance tests for the combined sample of boys and girls. As we can see, average differences in observable characteristics between treatment and control individuals are relatively

[24] This question was designed to explore gender bias in STEM conditional on the same math and science skills. See for instance Bertrand & Mullainathan (2004).

small and not statistically significant.[25] Note that the self-reported 10th grade GPA in all subjects, which we will use to measure pre-treatment academic aptitude, is very close among treatment and control students. It can be pointed out that the treatment may have influenced students incentives to either reveal or conceal their 10th grade GPAs in the follow-up survey; so the fact that almost no differences are observed alleviates such concerns. Regarding other individual characteristics, on average students are 16 years old, have 2 siblings, their parents have 13 years of education and in 85% of the famlies own their houses. About 95% of students have a working father and 68% have a working mother; while 15% and 3% have an engineer father and an engineer mother, respectively. Finally, the number of boys and girls in each quartile is also balanced across treatment and control schools.

Panel B in Table 1 compares the school level characteristics and finds that treated and control schools are similar on average. It is important to highlight that treated and control schools had almost the same performance in the 2015 national standardized evaluation for 8th graders; which corresponds to the year in which students in our cohort were evaluated. These results also support the use of self-reported GPAs as a reliable measure of students pre-treatment academic aptitude. Column 4 of Table 1 shows the p-value associated with the coefficient on the treatment indicator of a regression of each covariate on treatment status. The regression controls for city fixed effects, and standard errors are adjusted for clustering at the unit of randomization (school).

Gender differences in preferences, perceptions, beliefs and stereotypes: Here we focus exclusively on individuals in the control group to asses gender differences in terms of preferences, self-confidence, perceptions, beliefs and stereotypes. As shown in Table 5, while approximately 40% of boys in the control group preferred an engineering major, the number was only 14% for girls. The gap remains similar if we only focus on students in the top 10th grade math GPA quartile. In this case, more than 50% of boys stated engineering as their preferred college major, while just 20% of females did so (see Figure 2). The observed gap is

[25] F-stat for joint significance is 1.11 (p-value is 0.358) and hence we can reject that all the variables can jointly explain the assignment to treatment. Nonetheless, in our estimations we will also control for baseline characteristics to improve the precision of the estimates.

likely connected to the fact that girls are less confident than boys in their skills and capabilities to pursue a career in engineering, (37% versus 59%), as it can be observed in Table 5.

Both boys and girls are more likely to consider a successful professional engineer to be a male. The percentage is nevertheless higher for boys (88% for boys and 61% for girls). Interestingly, nearly half of boys and girls (52% and 49% respectively) suggested engineering as a college major to "Lorena": a hypothetical female high school friend with the highest possible math and science GPA scores. Among girls in the top quartile of the math GPA distribution, this percentage is 58%, which is in clear contrast to the low proportion of them that prefer engineering majors. This suggests that while high math ability girls are likely to project another high math ability girl into an engineer career, they are less likely to do the same for themselves.[26] These findings strongly point to the role of interventions (such as the role models we study in this paper) that can boost self-confidence of female students, especially those who possess the aptitude, to pursue engineering majors.

3.2 Empirical Strategy

We estimate the following Linear Probability Model (LPM):

$$Outcome_{isc} = \beta_0 + \beta_1 T_{isc} + \beta_2 female + \beta_3 female * T_{isc} + \beta_4 X_{isc} + \theta_c + \varepsilon_{isc} \quad (1)$$

where $Outcome_{isc}$ denotes the outcome of student i in school s and city c; T_{isc} is a dummy variable indicating whether the student's school located in city c has been selected to receive a role model visit, $female$ is a dummy variable that equals one for girls and zero for boys. We interact the female indicator with the treatment dummy to test for heterogenous treatment effects. We also control for student characteristics X_{isc} (including household background) and add city fixed effects (θ_c) to account for the fact that the randomization was stratified by city. Finally, in all our estimations standard errors are clustered at the school level.

The estimate on T_{isc} captures the Intent-to-Treat (ITT) effect of our

[26]Note in Table 5 that students, both male and female, on average are able to list 4 types of engineering majors.

intervention, since compliance with the initial random assignment was not perfect.[27] To deal with the non-compliance, we also estimate the local average treatment effect (LATE) using random assignment as an instrument for actual treatment. The LATE estimates are very close to the ITT ones and are shown in the Appendix, Table A2.

A possible concern is that treated students may have talked about the role models talks contents with peers in control schools (i.e. friends in the neighborhood who attend a different school, or siblings attending different schools), which we expect to happen with low probability since the school or even the class is the unit within most peer interactions take place [Avvisati et al. (2014)]. Nevertheless, if spillovers do exist, our estimates could be interpreted as a lower bound of the actual impact of the intervention on students' career preferences.

4 Results

4.1 Effects on Preferences for Engineering Majors

Table 2 presents the ITT intervention effect on students' preferences for engineering majors following the specification in equation (1). The first column shows the effect on students' preferences without controlling for covariates. Control variables are added gradually in columns 2 to 6. Notice that controlling for covariates does not significantly change either the sign or the size of the estimates. For the overall sample, the intervention does not have a statistically significant impact on boys' and girls' preferences for engineering programs.[28]

Regarding other covariates included in columns 2 to 6 in Table 2, several patterns are worth mentioning (See Table A1 in the Appendix). Firstly, in Peru engineering is clearly a male domain, and girls are 26 percentage points less likely to prefer engineering majors than boys (significant at 1%). Within-household peer effects are also likely to be present. Students with

[27] Close to 12% of the schools assigned to the treatment group could not be visited by the role models. Non-compliance was mostly related to schools administrators not allowing the visit to take place as well as last minute cancellations due to other school activities taking place.

[28] Also, Figure A7 in the Appendix shows the intent to treat estimates and the effect of other covariates on senior-year students' major preferences, while Figure A8 in the Appendix reports the ITT estimates and the effect of other covariates on students' perceptions.

female siblings who are engineering students are 6 percentage points more likely to prefer engineering majors (significant at 5%). Similarly, those with a father engineer are 4 percentage points more like to state such preferences. Wealth also plays a role, and students who own their houses are more likely to prefer engineering majors by 3 percentage points. Finally, our results also point to comparative advantage in skills as a factor that is strongly related to preferences for majors. An additional point in grade 10th math GPA is related, *ceteris paribus*, to a 5 percentage points (significant at 1%) increase in the likelihood of preferring an engineering major. Similarly, an additional point in Language (Spanish) 10th grade GPA relates to a 2 percentage points (significant at 1%) decrease in the likelihood of preferring engineering as a major of study.

Given the recent findings in the role models and STEM career choices literature,[29] next we explore if our intervention had heterogeneous effects for different ranges of the students' math ability distribution, as measured by their self-reported 10th grade math GPA. Then we look for local role model effects. That is, if role models effects are stronger among students who reside closer to UDEP historical area of influence.

Heterogeneous effects as a function of math ability. To shed light on how our role models intervention might have impacted students differently depending on their math aptitude, we split the sample into four groups or quartiles as a function of their math GPA in the school year preceding the intervention (10th grade).[30] As shown in Table 1 Panel A, the number of girls and boys in each quartile is balanced across treatment and control groups.[31]

Figure 2 shows the proportion of senior-year high school students who listed engineering as their most preferred college major, separately by gender and over the quartiles of pre-treatment math GPA. Not surprisingly,

[29] Several papers have explored heterogeneous effects of exposure to role models on educational outcomes such as Kipchumba et al. (2021), Lim & Meer (2020), Porter & Serra (2020), and Breda et al. (2020).

[30] These four groups or quartiles are constructed based on the students self-reported 10th grade math GPA. Considering a 20-point grading scale, students in the fourth, highest, quartile have a baseline math GPA in 10th grade higher than 16, those in the third quartile have baseline math scores of 16, those in the second lowest quartile have baseline math scores of 14 or 15, and finally those in the first, lowest, quartile report baseline math scores less than or equal to 13. Moreover, since baseline self-reported GPA math scores are discrete, the quartiles constructed do not have similar sizes: 32% of the observations lay in the first or lowest quartile, 34% of the observations lay in the second quartile, 14% of the observations lay in the third quartile, and 20% of the observations lay in the upper or top quartile.

[31] With the exception of boys in Q3 which is slightly higher in treatment relative to the control group.

students in the top math GPA quartiles find engineering majors more attractive. We can also observe that our intervention seems to have a positive impact only on girls in the top quartile of the math GPA distribution. For this particular subgroup, the probability of preferring engineering as a college major increases by 7.3 percentage points (significant at 10%) if their school was assigned to the role model intervention. To put this in perspective: this result represents a 36 percent increase from the 20 percent baseline level and amounts to a 18.6 percent reduction in the gender gap. There is no evidence in Figure 2 of any intervention effect among boys in the upper math GPA quartiles. Again this is due to the fact that these boys are already strongly committed to engineering majors: more than 50% of them indicated preferences for an engineering field. In social and cultural contexts in which high math ability boys are already committed to pursuing an engineering major, a soft role model intervention that mainly targets females is unlikely to have an impact on their preferences.

Table 3 explores the ITT intervention effects for different subsamples of students based on math ability. The ITT estimated coefficient for females in the top math quartile is positive and statistically significant at the 5% level; while the estimated ITT coefficient among boys in the upper math quartile is close to zero and not statistically significant. Note however that the coefficients for the interaction term among treatment and female status are not statistically significant; so we cannot reject the null hypothesis of no different ITT effects across genders.

While Figure 2 clearly indicates that boys in the upper math quartiles are not affected by the intervention; interestingly there seems to be a positive, although not statistically significant, effect among boys in the lower quartiles. To explore this further, columns (1) and (2) in Table 3 focus on the two lowest math GPA quartiles. While the ITT male estimates are always positive, they are not statistically significant.

Finally, there may be some concerns given our use of self-reported math scores as a proxy for ability. In particular, the intervention itself may have influenced how students reported their baseline math scores. For example, social desirability bias may have led students in the treatment group to inflate their scores relative to those in control schools. As we discussed before, there are no statistically significant differences in terms of self-reported 10th grade among treated and control individuals. Nevertheless, to examine this further, we compare the distribution of reported math scores (from the survey) to actual math scores using administrative data.[32] As can be seen in Figure A9 and Table A12 in the Appendix, students tend to inflate their math scores. The average school inflation is 1.42 points. However, there is no difference in over-reporting between treatment and control schools (1.43 for treatment and 1.40 for control). Further, we find that girls inflate less than boys. It is possible that this difference in grade inflation affects the composition of quartiles. However, we find that our results remain unchanged if we use gender-specific distributions to create quartiles.[33]

Heterogeneous effects as a function of geographical location. During their talks, role models clearly stated their UDEP connection. In this sense, they may have been more effective (i.e. better at capturing the student attention) in schools within areas in which our partner university has a relatively high reputation and/or recognition. In fact, UDEP is recognized as the most prestigious university in the Piura region, and this influence and prestige also extends to the neighboring regions of Tumbes (to the North) and Lambayeque (to the South). This is confirmed by the fact that in the last 5 years, 80% to 85% of incoming UDEP's students are from Piura, and close to 95% are from the three above-mentioned regions. On the other hand, the regions of La Libertad, Cajamarca, and Ancash, which are geographically distant from Piura, have their own established local and regional universities.[34]

In Figure 3 we restrict our sample to schools within the Piura region and observe that the ITT effect for girls in the top math quartile becomes stronger (15.7 percentage points) and highly statistically significant at the 1% level. Taking this evidence into account, in Table 4 we explore the ITT intervention effects in schools located in Piura and the neighbouring regions

[32] All schools submit Grade 10 scores to the Ministry of Education through SIAGIE platform, 'Sistema de Información de Apoyo a la Gestión de la Institución Educativa'. We have access to this data and restricted it to schools in our sample to construct the actual distribution of math scores. However, the student identifier in this data set is anonymized because of which we are unable to use students' actual math scores in our estimation.

[33] Results available upon request.

[34] For example, in La Libertad region, Universidad Privada del Norte, Universidad Antenor Orrego, and Universidad Nacional de Trujillo are generally identified as the three most prestigious regional universities.

of Tumbes and Lambayeque. Our findings indicate that our intervention seems to have been more effective at steering high math skilled girls (upper quartile) towards engineering fields in schools located geographically close to UDEP. Treated girls in the top math GPA quartile and enrolled in schools at neighbouring regions are 13.1 percentage points more likely to prefer engineering (77% increase from a baseline of 17%) than similar girls in the control group. Note that the interaction coefficient term between the treatment and female indicators is statistically significant at the 1% level. Hence, for the Piura, Tumbes, and Lambayeque regions altogether, we can clearly reject the null hypothesis of no difference in treatment effects between boys and girls in the top math quartile. In effect this suggests that the role models are effective at persuading only those students who know about UDEP and its programs.

In column (1) and (2) of Table 4, we explore the ITT effects among students in the two lowest math quartiles, for the Piura, Lambayeque and Tumbes regions altogether. The ITT among low ability boys is always positive (between 6 and 7 percentage points) but not statistically significant. Note that the estimated effect for girls in this case is very close to zero and not statistically significant. Since role models emphasized that a person does not have to be a math genius to major in engineering, and that skills like imagination and creativity are also important to pursue engineering careers, it seems possible that some low math ability boys may have adjusted their engineering preferences as a result of this specific message. Nevertheless, given the weak nature of the evidence in this case, we believe these results should be treated with caution.

Table A17 in the Appendix summarizes the ITT estimates for our role model interventions separately by gender, geographical location and 10th grade math GPA quartile. In Panel A of Table A17, the ITT estimates correspond to the full sample, while in panel B we restrict the analysis to schools in the region of Piura and adjacent regions of Tumbes and Lambayeque. Overall, our intervention increased the likelihood that a female senior-year high school student in the upper math quartile who resides geographically close to UDEP, stated engineering as her most preferred major. No effects for female students in the 1st, 2nd, and 3rd quartile were found in any specification.[35] Moreover, we also find some, although week, evidence suggesting that female role models may have also affected the engineering major preferences of low math ability male students.

4.2 Self-confidence, Beliefs and Stereotypes

To shed light on the mechanisms behind our intervention, in the follow-up survey we presented students with several statements intended to measure self-confidence and gender stereotypes. We begin this analysis with Table 6, which focuses on students' self-confidence in their own skills and aptitude to pursue engineering majors. We observe that treated girls in the top quartile of the math score distribution in Piura, Tumbes, and Lambayeque schools are 12.5 percentage points (significant at 5%) more likely to indicate that they do have the necessary skills and aptitude to major in engineering. This specific result suggests that role models were a source of credible information and inspiration, positively influencing these girls' self-confidence, which, as shown before, resulted in an increased preference for studying engineering.

Interestingly, Table 6 also shows that high math ability treated boys appear to be less confident in their aptitude and skills to pursue an engineering major. As pointed before, one of the key pieces of information in the role models talks was that you don't need to be a mathematical genius to major in engineering, and that other skills, such as imagination and ingenuity, are also relevant. This message may have influenced the perceptions of these boys about the role of math skills alone to succeed in engineering majors. Note however that the adjustment in perceptions did not affect their stated preferences. In a social context in which high math ability boys are expected to be engineers, this extra piece of information and the subsequent adjustment in perceptions, is likely not to be adequate to fully switch them out of engineering majors. In any case, such issues should be kept in mind when designing similar role model interventions. Also interestingly, in the case of boys in the two lowest math ability quartiles, the results in Table 6 suggest a 4 to 5 percentage points increase in the

[35] We also calculated probit marginal effects, which are similar to the OLS estimates and can be provided upon request.

self-confidence outcome; however, it is not statistically significant. Once again, the key message on aptitudes other than math ability to succeed as an engineer may be playing a role in this case. Overall, boys within the UDEP's catchment area seem to have been carefully listening to the information provided in the talks.

In Tables 7 and 8, we evaluate whether or not the role models affected gender beliefs, biases and stereotypes. In our follow-up survey, we described a person who happens to be a successful engineer and asked students whether they thought that this person was more likely to be male or female. We constructed an indicator that took the value of one when the student responded that the person was more likely to be male, and zero otherwise. Table 7 presents the results related to this question. In general, the coefficients for all females quartiles have the expected sign: treated girls are less likely to indicate that the successful engineer is male, but the estimated coefficients are relatively small and not statistically significant. The estimated coefficients for low ability boys are also negative, but smaller in absolute size than the female ones and not statically significant. Again this suggests that a lasting impact on gender beliefs, biases and stereotypes, possibly needs a longer than 20 minutes intervention in countries like Peru.

Also regarding gender stereotypes, we presented students in our sample with two hypothetical high school students: a female named "Lorena" and a male named "Javier", and describe both as high math and high science ability individuals. We then asked students to suggest a major to each of them. The outcome variable takes the value of one if the student recommended an engineering major to "Lorena". The results in Table 8 indicate that treated girls were not more likely than control ones to recommend engineering majors to "Lorena". It is nevertheless important to note that close to 60% of top math ability girls in the control group are already recommending engineering to the hypothetical high math ability girl. However, in the treatment schools very few of these high math ability are applying this recommendation to themselves. In other words, self-confidence seems to be the critical issue; and as we have shown before, it is self-confidence what is primarily being impacted by our role models. In the case of treated boys, column (2) in Table 8 indicates that those in the second-lowest quartile residing within UDEP catchment area were more likely to suggest an engineering major to our hypothetical female student. As mentioned before, the fact that role models strongly emphasized that women can also succeed as engineers may be the driving factor behind the boys results. Clearly boys seem to have been paying attention to the information delivered.

Although the role models did not provide either any information about earnings associated with engineering careers (as they wanted it to be about abilities) or an exhaustive list of engineering specializations, in the follow-up survey we asked students related questions as we wanted to see if this intervention made them seek out more information on engineering majors. In this case, we find some statistically significant effects for girls in the top math quartile and boys in the second lowest math quartile; however, the size of the effects is relatively small. As we can see in Tables 9 and 10, treated high math ability girls listed 0.2 less engineering fields than those in the control group (who listed 4.7); while the salary expectations of boys in the second lowest math quartile increased just by 1% relative to the control group.

4.3 Effects on Types of Engineering

Our role models were from one of the following three engineering majors: industrial, civil and mechanical-electrical, which are actually the only three engineering specializations offered at UDEP. During their presentations, the role models emphasized their own major as well as their connection with UDEP. Given this context, role models may have been more effective at promoting their own engineering major or UDEP engineering majors in general. To test for this possibility, we create a binary outcome variable which equals one if the student stated as her/his preferred engineering major to be any of the role models' ones and zero otherwise. As shown in Table 11, girls in the top math ability quartile and within UDEP's catchment area are 13.4 percentage points (significant at 1%) more likely to list one of the role models' engineering majors. Note also that there is no statistically significant effect among boys, and the estimated coefficients in this case are relatively small. These results suggest that treated female students are clearly connected with the specific experience of their role models, and confirm that role models were a source of inspiration and were successful at boosting their self confidence. The results also provide important lessons for the design of role model interventions. STEM role models seem to be more effective at influencing career paths that are closely related to their own experiences. In this regard, it is important to emphasize that while UDEP is the leading university in the area, there are several other universities, including the public National University of Piura, which offer a wide range of engineering majors in addition to UDEP's ones. Hence, the students' major preferences are unlikely to be fully restricted by UDEP engineering academic offer. [36]

4.4 More on Local Effects: Proximity to UDEP

In this section we provide additional evidence on whether students in schools located geographically close to UDEP were more likely to be encouraged to pursue engineering fields by our role models, but also on whether or not the effects in UDEP proximate schools are different from those in far-away ones. Using longitude and latitude coordinates, we calculate the distance in kilometers from the schools in our sample to UDEP. Using the estimated distances, Tables A5-A10 in the Appendix show that geographical closeness matters. Girls in the top GPA math quartile are 17.2 percentage points (significant at 1%) more likely to prefer engineering after a role model exposure if they come from a school located below the median distance (less than 43 km) from UDEP (Column 6, Table A5 in the Appendix). Moreover, the effect for schools above the median distance is close to zero and we can reject the null hypothesis of equal ITT effects among nearby and far-away schools. The analysis for boys in the bottom quartiles of math scores is presented in Table A6 in the Appendix and leads to similar conclusions.[37] Overall, these results confirm that the relevance of UDEP's role models and their ability to inspire is higher in the geographical areas where UDEP historically has had a stronger influence.

[36] It could be possible that the significant effect on preferences for the role models' engineering types could be due to having more students who prefer those academic programs. 81% of students that preferred engineering stated preferences for engineering programs of role models (75% girls and 84% boys). Overall, 21% of students in the follow-up survey preferred the role models' engineering programs.

[37] In Tables A7-A10 in the Appendix we evaluate the effect on different subgroups of individuals based on location and math skills. The results seem to be robust for different subgroups of students. The liking for engineering increases after the intervention in nearby schools.

4.5 Effects on Other Majors

In the previous section, we found that, relative to the control group, treated high math ability female students increased their preferences for engineering. In this section we explore how the intervention affected students' preference for both non-STEM majors and STEM majors other than engineering.

As we can observe in Table A25 in the Appendix, the intervention clearly affected the preferences for non-STEM majors among high math aptitude girls in UDEP proximate schools. Treated girls in the top math quartile are less likely to report they will choose a non-STEM major.[38]Also note that the absolute value of the estimated coefficients are relatively similar to those observed for high math aptitude girls in Table 4. This clearly indicates that our intervention is shifting the preferences of high math aptitude girls from non-stem majors to engineering ones.[39] There is also some, though weaker, evidence in this table suggesting that the intervention is doing the same for boys in the lowest math ability quartiles within UDEP's catchment area. In a similar fashion, we also investigate if students in our intervention were more likely to prefer STEM fields other than engineering (i.e. Life-Science, Mathematics, Statistics, Physics) as a consequence of being exposed to a young engineer role model. We do not find neither sizable nor statistically significant estimates in this case. This quite likely due to the extremely low share of students who prefer STEM majors different from engineering in Peru.[40]

5 Robustness Checks

5.1 Alternative measures of ability: Science and math

In our baseline estimations, we explored differential ITT effects as a function of students' math ability only, which is generally regarded an indicator of students' capacity to major in engineering. In this section,

[38]Non-STEM majors affected by the intervention are Business Administration, Economics, Communication, Accounting, Marketing, Law, Architecture, Medicine, Psychology.

[39]An alternative explanation of why we have exactly opposite effects of the intervention on preferences for non-STEM fields is the limited number of students who preferred a STEM field other than engineering.

[40]These results are available upon request.

in addition to math scores, we also consider 10th grade science scores, as competence in science may also be an indicator of the student aptitude to major in engineering. We therefore construct a binary indicator equal to one if the student ranked in the top quartile in both math and science, and zero, otherwise. As expected, the ITT effect in this case is positive and statistically significant for female students in the top, math and science, quartile and who reside within UDEP's catchment area (See Appendix Table A3). These girls are 21 percentage points (significant at 1%) more likely to prefer engineering as a result of our role models intervention.

5.2 ECE Math Scores

While our balance tests confirm that treatment and control schools are of similar academic quality, the fact that we are not using standardized scores to define our math ability quartiles may be a cause of concern for some readers. In order to alleviate these concerns, we again estimate the specifications in Table 4 using the school average math section performance in the *Evaluación Censal de Estudiantes* (ECE), which is a standardized national examination administered annually to 8th graders, as a control variable. As students in our sample were in the 8th grade in 2015, the 2015 ECE results allow us to control for the academic quality of our school cohorts. The estimations results are shown in Table A4 in the Appendix.

5.3 Missing Information and Attendance Rates

The follow-up survey was administered to 11th graders in November 2018 in both treatment and control schools. Respondents to the survey were students who attended schools the day of the survey administration. We investigate the effect of students' selection to attend school the day of the survey on our main outcomes of interest. The results show that not all students attended schools when the survey implementation took place (but attendance rates were high 90%), however, there is no statistically significant difference in attendance rates between treatment and control groups. Attendance rate in treatment schools was 95% and in control schools was 87%. A regression analysis shows similar results and indicates that the coefficient on treatment is not statistically significant. Table A13 in the Appendix shows the estimates and attendance school rates for students in treatment and control schools and separately by girls and boys.

Furthermore we investigate if there exists differences in non-reporting answers to some of the questions in the follow-up survey between treatment and control schools. We might be concerned of bias introduced in the estimation due to self-selection to report. According to the results depicted in Table A14, we are not very concerned about missing values in the follow-up survey questions as we cannot reject the null hypothesis of same non-reporting behavior between treatment and control schools. Non-reporting behavior was measured by the number of students in a given school who did not answer at least one of the questions related to preferences for engineering, STEM fields, self-confidence about own skills to succeed in engineering and gender stereotypes among others.

5.4 Multiple Hypothesis Testing

Our study's purpose is to evaluate the effect of a role model interaction on students' preferences for engineering programs. Thus, we only have one single primary outcome, 'preference for engineering fields' in our study. Suppose we are interested in testing the effect of our role model treatment on a range of multiple outcomes corresponding to students' career preferences and perceptions regarding women in engineering fields. Let's define our outcomes: i) student's preference for any engineering field, ii) student's preference for a type of engineering selected by role models (i.e. industrial engineering, civil engineering, or mechanical engineering), iii) student's preference for non-STEM programs, iv) students' self-confidence in own math ability to succeed in engineering fields, v) students' knowledge about engineering types, vi) students' knowledge about salary in the engineering profession, vii) gender stereotypes in STEM, and viii) students' beliefs about gender success gaps in STEM fields.

We run the following treatment regression for outcome j:

$$Y(j) = a + b1 * Treatment + c1 * y(j,0) + d'X + e(j) \tag{2}$$

Equation (2) follows an Ancova specification and controls for baseline values of the outcome variables included in y(j,0) term. X is a matrix of covariates including the randomization strata (city fixed effects). e is the residual.

Testing for the effect of a treatment in multiple outcomes separately produces false rejections, and we might erroneously conclude that the treatment has a significant effect on outcomes when it is not the case.

Recall that the probability to reject the null hypothesis when it is true is a false rejection. Using a critical value of 0.05, the false rejection rate with 8 p-values is $1 - 0.95^8 = 0.34$. To reduce the probability of false rejections, previous studies have used multiple hypothesis testing. For instance, Anderson (2008) computes sharpened false discovery rates, q-values, adjusting p-values for multiple test hypothesis.[41] Table A15 and Table A16 in the Appendix, show p-values and q-values associated with the ITT estimate for high ability girls and low ability boys in local schools, respectively. We explore the effect on the main outcomes of interest such as preference for fields of study and perceptions regarding STEM fields. Adjusting for multiple hypothesis, the coefficient on preference for engineering, preference for role models' major of study, preference for non-STEM fields, and students' confidence in own abilities to succeed in engineering are statistically significant at 5% for high ability girls. By contrast, the treatment did not have statistically significant effects on low-ability boys, neither on career preferences nor perceptions.

6 Conclusion

Using experimental evidence from an RCT in Peru that exposed senior high school students to young female engineers (college seniors or recent graduates), we show that role models are important as they are able to influence preferences for some students. We find that as a result of the treatment, girls in the highest math ability quartile are more likely to prefer engineering majors, and weaker evidence for boys in the two lowest math quartiles. We show that the role models were able to inspire these girls by changing their self-confidence regarding their own skills and about the aptitudes necessary for successfully pursuing engineering majors. Thus,

[41] Anderson's code is available online.

high ability girls in the treatment schools revise their posterior beliefs about their own ability to succeed in engineering leading to a change in preferences. Interestingly, role models also affected the preferences of low math ability boys by conveying to them that engineering is not just about math ability; creativity and ingenuity are also important to succeed as engineers.

We also find that while role models matter, the context in which they intervene critically determines their effectiveness. In our study geographical proximity turns out to be important since the role model effects are stronger among students who attend schools located in the area where UDEP has historically had a stronger influence. Thus, the role model messaging and credibility is likely to be limited to a specific area of influence.

Importantly, we show that role models are more effective at influencing preferences for career paths that are closely related to their own experience, i.e., if we want girls to pursue civil engineering, the role models who have graduated from civil engineering programs will be more effective. This suggests that participants in the treatment schools were paying close attention to the role models. Moreover, high ability math girls were inspired by the information presented to them and were willing to change their career choices based on this. Both the geographical and/or disciplinary proximity ideas suggest that context that enables participants to identify with the roles plays a key role in their effectiveness.

All the above suggest that low-cost, "light touch" role model interventions can be effective in influencing gender gaps in STEM field participation. In particular, it can help by boosting the self-confidence of a small group at a low cost but will not be adequate for addressing deeper gender stereotypes. However, careful attention must be paid to design features in order to maximize their impact. Firstly, it is important to pay careful attention to their message content and potential audience. While female role model interventions related to STEM fields primarily target girls, boys may also react to some specifics of the message. Secondly, an important message for the design of role models programs is that the implementation context should be carefully evaluated. Not everyone can be an effective role model in every situation or at promoting any STEM field.

References

Anderson, M. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects', *Journal of the American Statistical Association* **103**(484), 1481–1495.

Ashraf, N., Bau, N., Low, C. & McGinn, K. (2020), 'Negotiating a better future: How interpersonal skills facilitate intergenerational investment', *The Quarterly Journal of Economics* **135**(2), 1095–1151.

Avvisati, F., Gurgand, M., Guyon, N. & Maurin, E. (2014), 'Getting parents involved: A field experiment in deprived schools', *Review of Economic Studies* **81**(1), 57–83.

Bayer, A. & Rouse, C. E. (2016), 'Diversity in the economics profession: A new attack on an old problem', *The Journal of Economic Perspectives* **30**(4), 221–242.

Beaman, L., Duflo, E., Pande, R. & Topalova, P. (2012), 'Female leadership raises aspirations and educational attainment for girls: A policy experiment in India', *Science* **335**(6068), 582–586.

Bear, J. B. & Woolley, A. W. (2011), 'The role of gender in team collaboration and performance', *Interdisciplinary Science Reviews* **36**(2), 146–153.

Bertrand, M. & Mullainathan, S. (2004), 'Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination', *The American Economic Review* **94**(4), 991–1013.

Bettinger, E. P. & Long, B. T. (2005), 'Do faculty serve as role models? The impact of instructor gender on female students', *American Economic Review* **95**(2), 152–157.

Blau, F. D. & Kahn, L. M. (2017), 'The gender wage gap: Extent, trends, and explanations', *Journal of Economic Literature* **55**(3), 789–865.

Bottia, M., E.Stearns, Mickelson, R., Moller, S. & L.Valentino (2015), 'Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM', *Economics of Education Review* **45**, 14–27.

Breda, T., Grenet, J., Monnet, M. & van Effenterre, C. (2020), 'Do female role models reduce the gender gap in science?: Evidence from French high schools', *IZA Institute of Labor Economics, Discussion Paper Series IZA DP No. 13163* .

Brooks, W., Donovan, K. & Johnson, T. (2018), 'Mentors or teachers? Microenterprise training in Kenya', *American Economic Journal: Applied Economics* **10**(4), 196–221.

Buser, T., Niederle, M. & Oosterbeek, H. (2014), 'Gender, competitiveness and career choices', *Quarterly Journal of Economics* **129**(3), 1409–1447.

Carrell, S., Page, M. & West, J. (2010), 'Sex and science: How professor gender perpetuates the gender gap', *The Quarterly Journal of Economics* **125**, 1101–1144.

Correll, S. J. (2001), 'Gender and the career choice process: The role of biased self-assessments', *American Journal of Sociology* **106**(6), 1691–1730.

Dee, T. S. (2007), 'Teachers and the gender gaps in student achievement', *Journal of Human Resources* **42**(3), 528–554.

Del Carpio, L. & Guadalupe, M. (2018), 'More women in tech? Evidence from a field experiment addressing social identity', *CEPR Discussion Paper DP13234* .

Eble, A. & Hu, F. (2020), 'Child beliefs, societal beliefs, and teacher-student identity match', *Economics of Education Review* **77**.

Eble, A. & Hu, F. (2022), 'Gendered beliefs about mathematics ability transmit across generations through children's peers', *Nature Human Behaviour* .

Ellison, G. & Swanson, A. (2010), 'The gender gap in secondary school mathematics at high achievement levels: Evidence from the American mathematics competitions', *Journal of Economic Perspectives* **24**(2), 109–128.

Emerson, T. L., McGoldrick, K. & Mumford, K. J. (2012), 'Women and the choice to study economics', *The Journal of Economic Education* **43**(4), 349–362.

Flory, J. A., Leibbrandt, A. & List, J. A. (2014), 'Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions', *The Review of Economic Studies* **82**(1), 122–155.

Gneezy, U., Niederle, M. & Rustichini, A. (2003), 'Performance in competitive environments: Gender differences', *Quarterly Journal of Economics* **118**(3), 1049–1074.

Hardt, D., Nagler, M. & Rincke, J. (2020), 'Can peer mentoring improve online teaching effectiveness? an rct during the covid-19 pandemic', *Cesifo 8671 Working paper* .

Hoffmann, F. & Philip, O. (2009), 'A professor like me: The influence of instructor gender on college achievement', *Journal of Human Resources* **44**(2), 479–492.

Hsieh, C., Hurst, E., Jones, C. & Klenow, P. (2019), 'The allocation of talent and U.S. economic growth', *Econometrica* **87**(5), 1439–1474.

Kahn, S. & Ginther, D. (2017), 'Women and STEM', *NBER Working Paper Series* .

Kipchumba, E. K., C.Porter, Serra, D. & Sulaiman, M. (2021), 'Influencing youths' aspirations and gender attitudes through role models: Evidence from Somali schools', *Working Paper* .

Kofoed, M. S. & McGovney, E. (2019), 'The effect of same-gender or same-race role models on occupation choice evidence from randomly assigned mentors at West Point', *Journal of Human Resources* **54**(2), 430–467.

Lafortune, J., Riutort, J. & Tessada, J. (2018), 'Role models or individual consulting: The impact of personalizing micro-entrepreneurship training', *American Economic Journal: Applied Economics* **10**(4), 222–245.

Lim, J. & Meer, J. (2017), 'The impact of teacher-student gender matches: Random assignment evidence from South Korea', *Journal of Human Resources* **52**(4), 979–997.

Lim, J. & Meer, J. (2019), 'Stereotypes, role models, and the formation of beliefs', *Journal of Human Resources* **54**(2), 430–467.

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

Macphee, D., Canetto, S. & Farro, S. (2013), 'Academic self-efficacy and performance of underrepresented stem majors: Gender, ethnic, and social class patterns', *Analyses of Social Issues and Public Policy* **13**(1), 347–369.

Neumark, D. & Gardecki, R. (1998), 'Women helping women? Role model and mentoring effects on female Ph.D. students in economics', *Journal of Human Resources* **33**(1), 220–246.

Niederle, M. & Vesterlund, L. (2007), 'Do women shy away from competition? Do men compete too much?', *Quarterly Journal of Economics* **122**(3), 1067–1101.

Niederle, M. & Vesterlund, L. (2010), 'Explaining the gender gap in math test scores: The role of competition', *Journal of Economic Perspectives* **24**(2), 129–144.

Nollenberger, N., Rodriguez-Planas, N. & Sevilla, A. (2016), 'The math gender gap: The role of culture', *American Economic Review* **106**(5), 257–261.

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

Reuben, E., Wiswall, M. & Zafar, B. (2017), 'Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender', *The Economic Journal* **127**(604), 2153–2186.

Rury, D. & Carrell, S. E. (2023), 'Knowing what it takes: The effect of information about returns to studying on study effort and achievement', *Economics of Education Review* **94**, 102400.

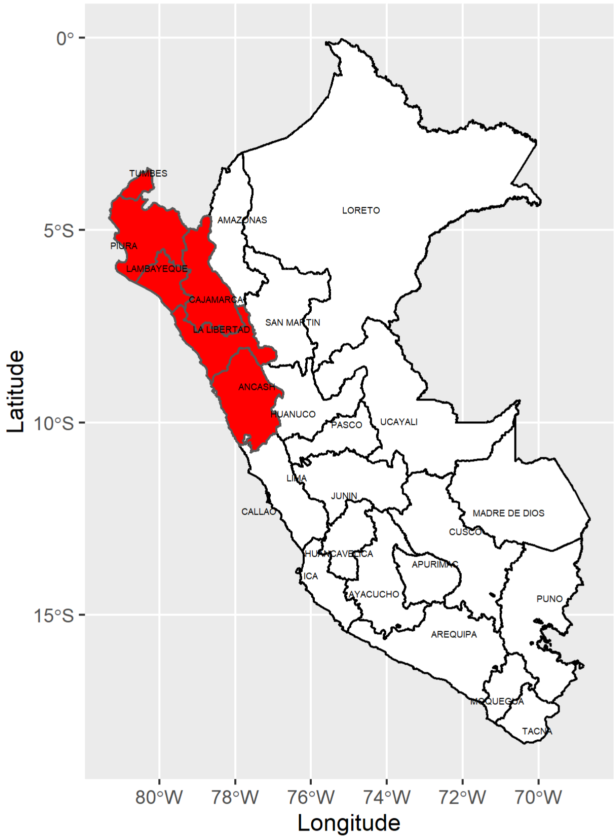
UNESCO (2017), 'Cracking the code: Girls and women's education in science, technology, engineering, and mathematics (STEM)', *Technical report* .

World Bank (2020), 'The equality equation, advancing the participation of women and girls in stem', *World Bank Technical Report* .

Zeltzer, D. (2020), 'Gender homophily in referral networks: Consequences for the medicare physician earnings gap', *American Economic Journal: Applied Economics* **12**(2), 169–197.

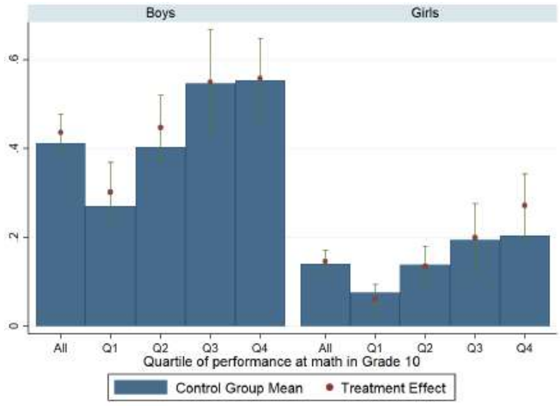
Figures and Tables

Figure 1: Experimental sample in Peru



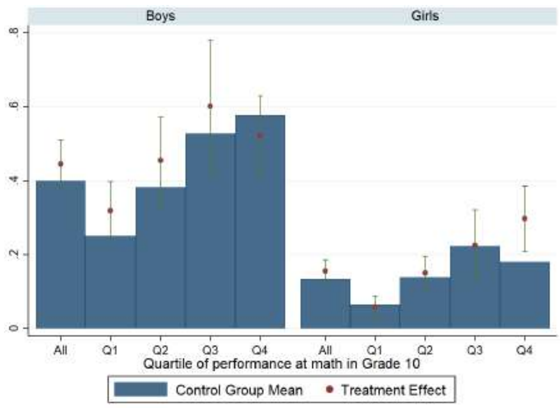
Notes: This figure shows the division of the Peruvian territory in 25 regions. The regions covered in our intervention are shaded red in the graph.

Figure 2: Senior-year high school students- preference for engineering by student gender and quartile of baseline math score



Notes: The figure shows the fraction of senior-year high school students (grade 11) who stated they would like to study Engineering after graduating from high school, for boys (left panel) and girls (right panel) separately. The blue bars indicate the mean among all students in the control group and the separate means by quartile of final course grade on math in grade 10. The red solid dots show the estimated treatment effects with 95% confidence intervals denoted by vertical capped bars.

Figure 3: Senior-year high school students- preference for engineering by student gender and quartile of baseline math score: only Piura



Notes: The figure shows the fraction of senior-year high school students (grade 11) who stated they would like to study Engineering after graduating from high school, for boys (left panel) and girls (right panel) separately. The sample includes only students in schools located in Piura. The blue bars indicate the mean among all students in the control group and the separate means by quartile of final course grade on math in grade 10. Red solid dots show the estimated treatment effects with 95% confidence intervals denoted by vertical capped bars.

Table 1: Treatment-control balance

	Control Group (1)	Treatment Group (2)	Difference T-C (3)	p-value (4)
Panel A: Student level (full sample)				
Female, gender (female=1)	0.575	0.540	-0.058	0.330
Age (in years)	16.232	16.266	0.018	0.393
Math, 10th grade math GPA	14.641	14.510	-0.083	0.621
Language, 10th grade spanish GPA	15.589	15.072	-0.333	0.100
Science, 10th grade science GPA	15.201	15.042	-0.170	0.278
Years education father	13.955	13.718	-0.185	0.279
Years education mother	13.641	13.419	-0.142	0.425
Father engineer	0.151	0.146	-0.014	0.411
Mother engineer	0.032	0.038	0.003	0.682
Number of siblings	1.959	1.962	-0.006	0.908
Own a house	0.845	0.854	0.009	0.508
Mother work	0.675	0.679	0.020	0.280
Father work	0.950	0.951	0.005	0.483
Has female sibling engineer	0.044	0.041	-0.003	0.599
(*)Girls in Q4 math	0.114	0.094	-0.013	0.529
(*)Girls in Q3 math	0.077	0.081	-0.002	0.921
(*)Girls in Q2 math	0.192	0.171	-0.030	0.351
(*)Girls in Q1 math	0.171	0.175	-0.009	0.719
(*)Boys in Q4 math	0.092	0.084	-0.001	0.914
(*)Boys in Q3 math	0.048	0.064	0.021	0.028
(*)Boys in Q2 math	0.140	0.149	0.011	0.605
(*)Boys in Q1 math	0.132	0.153	0.030	0.235
Number of Observations	2694	2704		
Test of joint significance excluding (*)	F-stat: 1.11 (p-value: 0.358)			
Panel B: School level (full sample)				
Average math ECE 2015	599.981	600.739	0.532	0.937
Number of teachers	14.944	16.660	1.518	0.457
Number of male teachers	7.882	9.136	1.251	0.367
Number of female teachers	7.500	8.106	0.451	0.743
Teachers-concluded pedagogy studies	23.755	27.326	3.320	0.383
Teachers-not concluded pedagogy studies	8.068	8.583	0.168	0.938
Private school	0.741	0.723	-0.051	0.555
Registration-total students	58.444	64.979	5.965	0.576
Registration-total male students	24.907	29.340	4.554	0.447
Registration-total female students	33.537	35.638	1.411	0.855
Single-sex school (only women)	0.130	0.128	-0.012	0.869
Test of joint significance	F-stat: 0.32 (p-value: 0.956)			

Notes: In panel A, the sample is restricted to students in the treatment and control groups who answered the post-treatment survey while in panel B the sample is restricted to schools in the treatment and control groups. Column 1 and column 2 report the sample mean in the control and treatment group, respectively. Column 3 displays the estimate on the treatment dummy in a regression of each variable on treatment. P-values for the statistically significance of the estimate are shown in column (4). The regression controls for city fixed effects, and standard errors are adjusted for clustering at the unit of randomization (school). A test for the joint significance of the coefficients is performed after running a regression of the treatment dummy on the baseline covariates. F-statistics are reported. Information in panel A comes from a follow-up survey implemented in 18 cities of Peru to senior-year high school students in November 2018 while in panel B the information comes from the *Censo Educativo 2017-MINEDU* and *Evaluación Censal de Estudiantes(ECE) 2015*.

Table 2: The effect of exposure to role models on students' preference for engineering

Dep. Variable:	Prefer Engineering					
Sample:	Full (1)	Full (2)	Full (3)	Full (4)	Full (5)	Full (6)
Treatment	0.036	0.036	0.035	0.034	0.016	0.018
	(0.025)	(0.026)	(0.026)	(0.024)	(0.024)	(0.024)
Female	-0.263***	-0.265***	-0.266***	-0.265***	-0.258***	-0.261***
	(0.018)	(0.018)	(0.018)	(0.017)	(0.019)	(0.018)
Interaction	-0.023	-0.024	-0.024	-0.024	-0.008	-0.008
(Treatment*female)	(0.027)	(0.028)	(0.028)	(0.027)	(0.027)	(0.027)
ITT female: Treatment + Interaction	0.013	0.011	0.011	0.010	0.008	0.010
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Number of observations (N)	5156	4872	4856	4783	4639	4580
Adjusted R^2	0.105	0.107	0.109	0.114	0.158	0.161
Mean Dv (Treatment==0)	0.14	0.14	0.14	0.14	0.14	0.14

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering for the full sample of students who answered the survey. Column 1 reports the ITT estimates without covariates. In columns 2 to 6 we gradually add the following controls: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 3: The effect of exposure to role models on students' preference for engineering (by quartile of math ability)

Dep. Variable:	Prefer Engineering			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	0.019	0.066	-0.069	-0.002
	(0.035)	(0.042)	(0.066)	(0.049)
Female	-0.195***	-0.254***	-0.338***	-0.307***
	(0.025)	(0.036)	(0.053)	(0.038)
Interaction	-0.039	-0.065	0.067	0.093
(Treatment*female)	(0.040)	(0.048)	(0.071)	(0.059)
ITT female: Treatment + Interaction	-0.019	0.001	-0.002	0.091**
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1437	1558	646	939
Adjusted R^2	0.118	0.146	0.147	0.136
Mean Dv (Treatment==0)	0.08	0.14	0.19	0.20

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering for students who answered the survey, separately by quartile of performance in math. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 4: The effect of exposure to role models on students' preference for engineering in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Prefer Engineering			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	0.058	0.064	-0.008	-0.043
	(0.036)	(0.048)	(0.072)	(0.050)
Female	-0.187***	-0.239***	-0.364***	-0.354***
	(0.025)	(0.041)	(0.062)	(0.042)
Interaction	-0.070*	-0.053	0.019	0.174***
(Treatment*female)	(0.041)	(0.054)	(0.084)	(0.065)
ITT female: Treatment + Interaction	-0.011	0.011	0.011	0.131***
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1132	1246	515	691
Adjusted R^2	0.131	0.135	0.150	0.141
Mean Dv (Treatment==0)	0.07	0.14	0.19	0.17

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering, separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 5: Difference in preferences for engineering and perceptions: by gender

Sample:	(1) Boys	(2) Girls	(3) Diff
Prefer engineering	0.405	0.139	0.266***
	(0.015)	(0.009)	(0.017)
Male_success	0.883	0.609	0.274***
Successful engineer is male	(0.010)	(0.013)	(0.016)
Self_confidence	0.585	0.367	0.219***
Consider to have needed skills to succeed in engineering	(0.015)	(0.012)	(0.019)
University_study	0.670	0.711	-0.041**
Plan to study at university	(0.014)	(0.012)	(0.018)
lorena_eng	0.520	0.492	0.028
Recommended engineering to Lorena	(0.015)	(0.013)	(0.020)
count_eng	4.323	4.403	-0.081**
Number of engineering majors listed	(0.031)	(0.023)	(0.037)

Notes: This table reports the means for different outcomes of a test of equality by gender. Standard errors are reported in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 6: The effect of exposure to role models on students' self-confidence in Piura/Lambayeque/Tumbes schools

Dep. Variable:	**Self-Confidence**			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	0.034	0.055	0.019	-0.116**
	(0.050)	(0.046)	(0.084)	(0.053)
Female	-0.199***	-0.232***	-0.208***	-0.294***
	(0.043)	(0.044)	(0.077)	(0.055)
Interaction	-0.030	-0.021	0.027	0.240***
(Treatment*female)	(0.056)	(0.056)	(0.101)	(0.084)
ITT female: Treatment + Interaction	0.003	0.034	0.046	0.125**
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1190	1290	522	708
Adjusted R^2	0.068	0.087	0.063	0.113
Mean Dv (Treatment==0)	0.19	0.34	0.51	0.55

Notes: This table reports the intent to treat (ITT) estimates on students' self-confidence in their aptitude and skills to pursue an engineering major, separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 7: The effect of exposure to role models on students' perceptions of males successfulness in engineering in Piura/Lambayeque/Tumbes schools

Dep. Variable:	**Males successfulness**			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	-0.026	-0.011	-0.007	0.014
	(0.036)	(0.032)	(0.063)	(0.048)
Female	-0.227***	-0.266***	-0.211***	-0.271***
	(0.043)	(0.041)	(0.065)	(0.048)
Interaction	-0.035	-0.034	-0.000	-0.023
(Treatment*female)	(0.060)	(0.047)	(0.077)	(0.065)
ITT female: Treatment + Interaction	-0.061	-0.045	-0.007	-0.009
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1126	1233	499	674
Adjusted R^2	0.083	0.107	0.057	0.093
Mean Dv (Treatment==0)	0.65	0.61	0.65	0.57

Notes: This table reports the intent to treat (ITT) estimates on students' perceptions of males successfulness in engineering, separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table 8: The effect of exposure to role models on students' recommending engineering to Lorena (hypothetical female friend) in Piura/Lambayeque/Tumbes schools

Dep. Variable:	**Engineering to Lorena**			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	-0.039	0.086*	-0.006	0.057
	(0.050)	(0.044)	(0.091)	(0.049)
Female	-0.035	-0.045	-0.048	0.084*
	(0.037)	(0.043)	(0.070)	(0.045)
Interaction	0.004	-0.082	0.046	-0.083
(Treatment*female)	(0.053)	(0.069)	(0.098)	(0.065)
ITT female: Treatment + Interaction	-0.035	0.003	0.039	-0.026
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1172	1270	520	697
Adjusted R^2	-0.005	0.021	0.004	0.025
Mean Dv (Treatment==0)	0.46	0.45	0.51	0.59

Notes: This table reports the intent to treat (ITT) estimates on students' recommending engineering to Lorena (hypothetical female friend), separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 9: The effect of exposure to role models on students' number of engineering fields listed in Piura/Lambayeque/Tumbes schools

Dep. Variable:	**Types of engineering listed**			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	-0.142	0.030	0.157	0.005
	(0.094)	(0.077)	(0.146)	(0.083)
Female	0.013	-0.057	0.004	0.139*
	(0.069)	(0.073)	(0.123)	(0.079)
Interaction	0.106	-0.034	-0.139	-0.208*
(Treatment*female)	(0.120)	(0.111)	(0.156)	(0.110)
ITT female: Treatment + Interaction	-0.035	-0.004	0.018	-0.203**
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1198	1296	525	710
Adjusted R^2	0.074	0.112	0.031	0.045
Mean Dv (Treatment==0)	4.32	4.33	4.51	4.65

Notes: This table reports the intent to treat (ITT) estimates on students' number of engineering fields listed, separately by quartile of performance in math.The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 10: The effect of exposure to role models on students' earnings expectations in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Salary (in logarithm)			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	0.022	0.095**	-0.090	-0.078
	(0.050)	(0.040)	(0.068)	(0.052)
Female	0.033	0.012	-0.103**	-0.043
	(0.050)	(0.038)	(0.048)	(0.049)
Interaction	-0.085	-0.094*	0.082	0.085
(Treatment*female)	(0.066)	(0.049)	(0.073)	(0.073)
ITT female: Treatment + Interaction	-0.063	0.002	-0.008	0.007
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1187	1291	523	707
Adjusted R^2	0.005	0.011	-0.002	0.001
Mean Dv (Treatment==0)	8.20	8.19	8.17	8.23

Notes: This table reports the intent to treat (ITT) estimates on students' knowledge about earnings associated with engineering careers, separately by quartile of performance in math.The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table 11: The effect of exposure to role models on students' preference for role models' engineering majors in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Any three types of engineering			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	0.051	0.052	-0.010	-0.073
	(0.038)	(0.042)	(0.073)	(0.047)
Female	-0.152***	-0.200***	-0.351***	-0.364***
	(0.026)	(0.035)	(0.056)	(0.043)
Interaction	-0.051	-0.029	0.028	0.208***
(Treatment*female)	(0.041)	(0.049)	(0.080)	(0.059)
ITT female: Treatment + Interaction	0.000	0.023	0.017	0.134***
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1132	1246	515	691
Adjusted R^2	0.124	0.104	0.151	0.143
Mean Dv (Treatment==0)	0.03	0.10	0.15	0.13

Notes: This table reports the intent to treat (ITT) estimates on students' preferences for the role models' majors offered at UDEP (industrial engineering, civil engineering, or mechanical engineering), separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Online Appendix

Figures in Appendix

Tables in Appendix

Figure A1: Program evaluation timeline

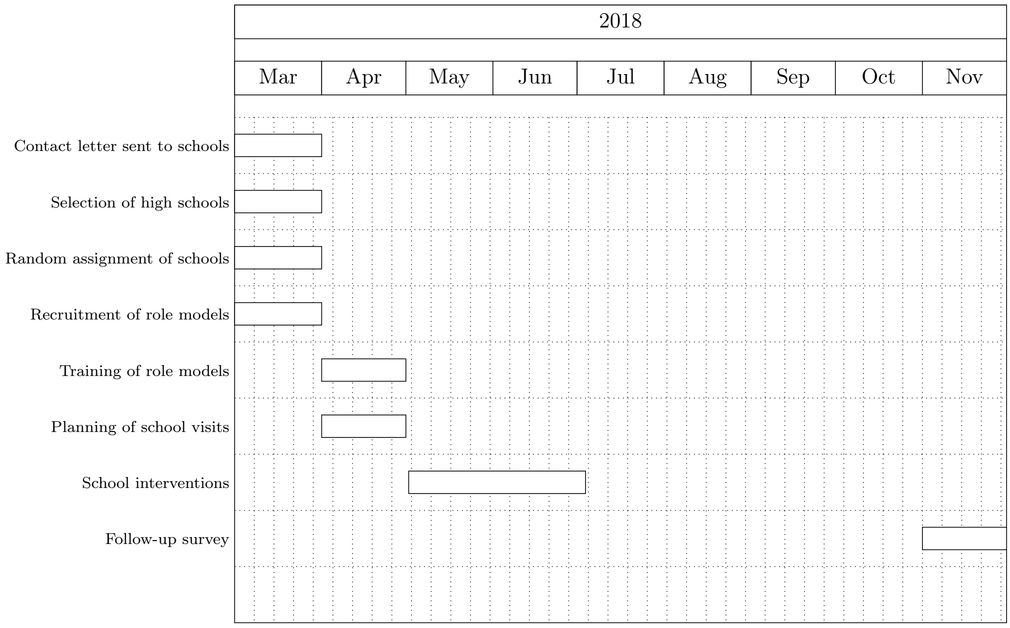
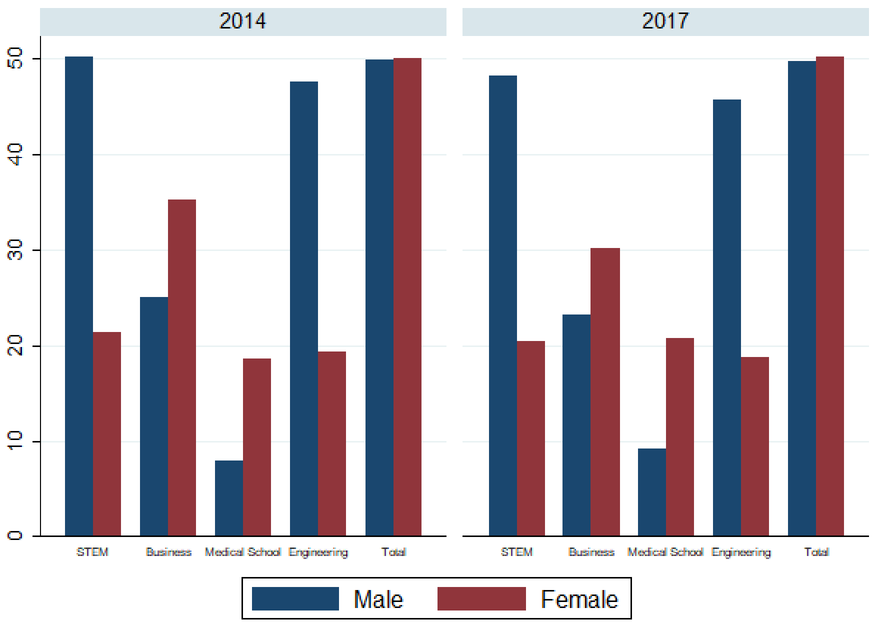


Figure A2: Thumbnails of slides shown during school visits

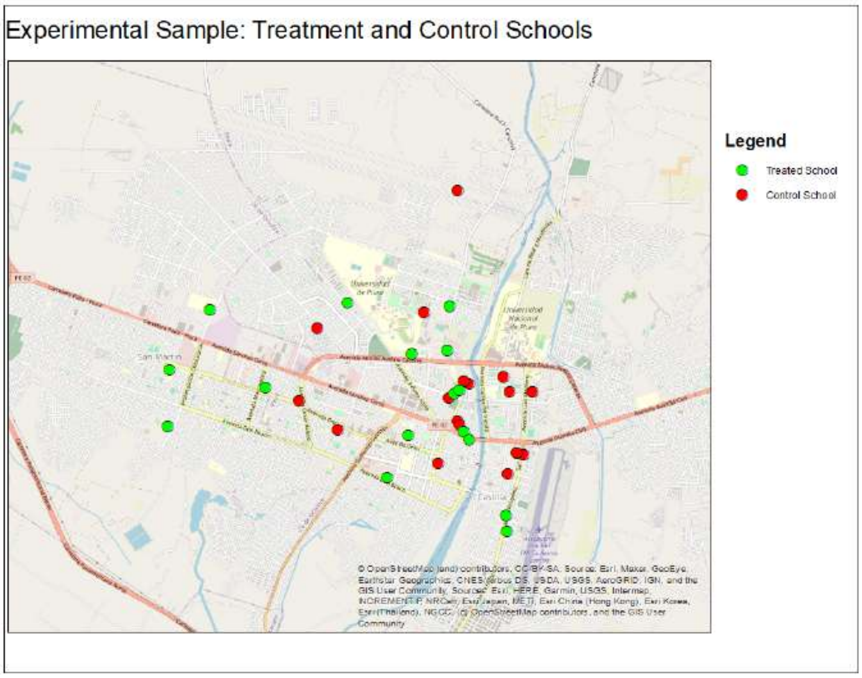
Figure A3: Thumbnails of slides shown during school visits (continued)

Figure A4: Share of male and female applicants to selective undergraduate academic programs for the whole population of applicants in 2014 and 2017, Peru



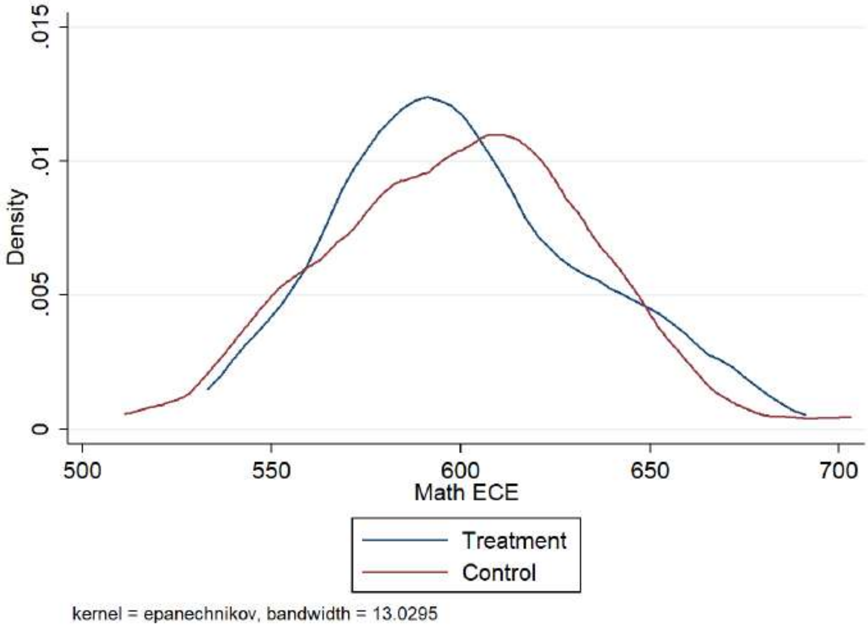
Notes: Data from public records of the Peruvian National Superintendence of Higher Education (SUNEDU): https://www.sunedu.gob.pe/sibe/.

Figure A5: Experimental sample: Piura city



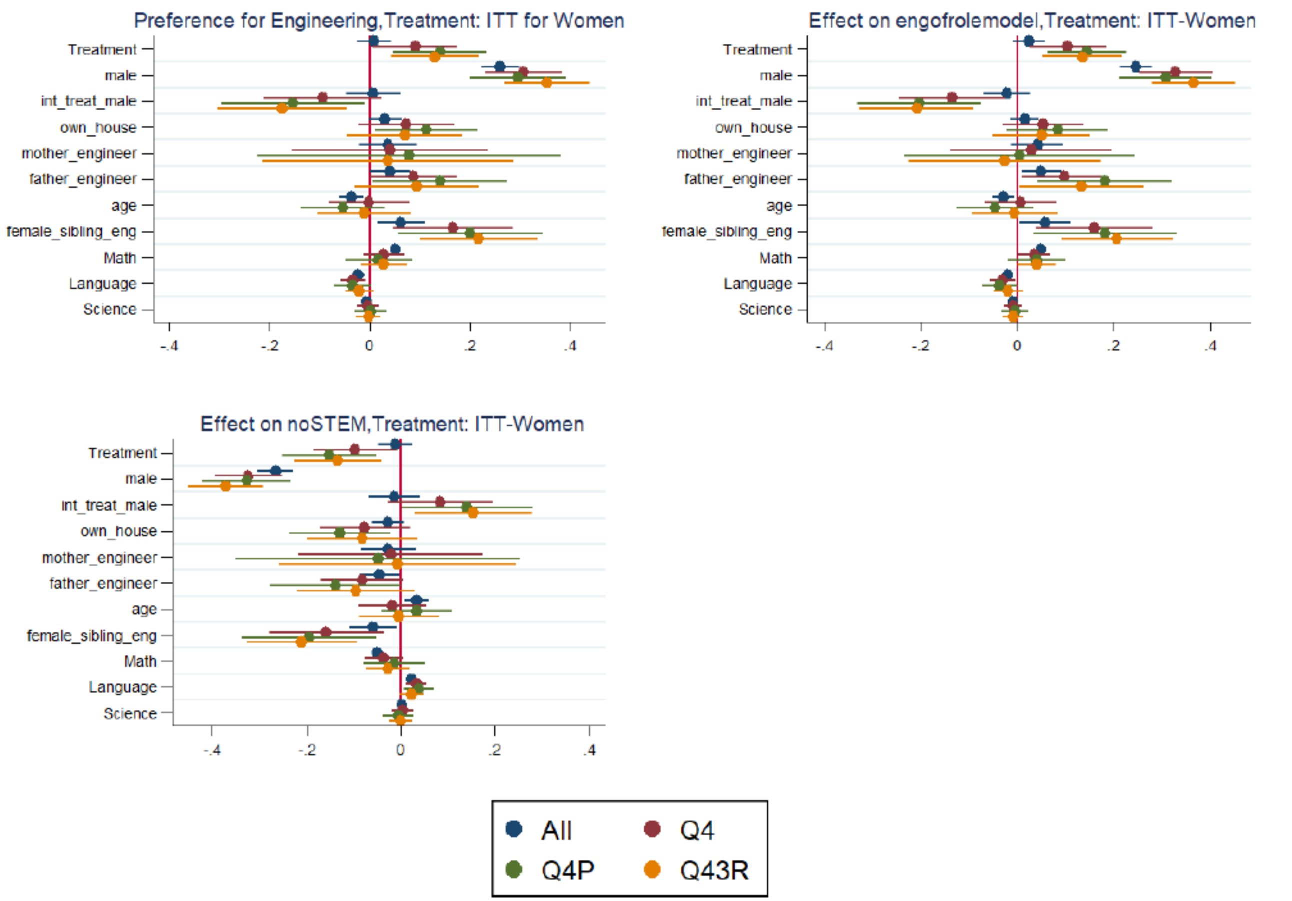
Notes: This figure shows the longitude and latitude coordinates of the schools in our sample. The sample is restricted to schools in Piura City (the role models' place of residence). The location of treatment and control schools are depicted with green and red dots, respectively.

Figure A6: Distribution of school ECE math scores



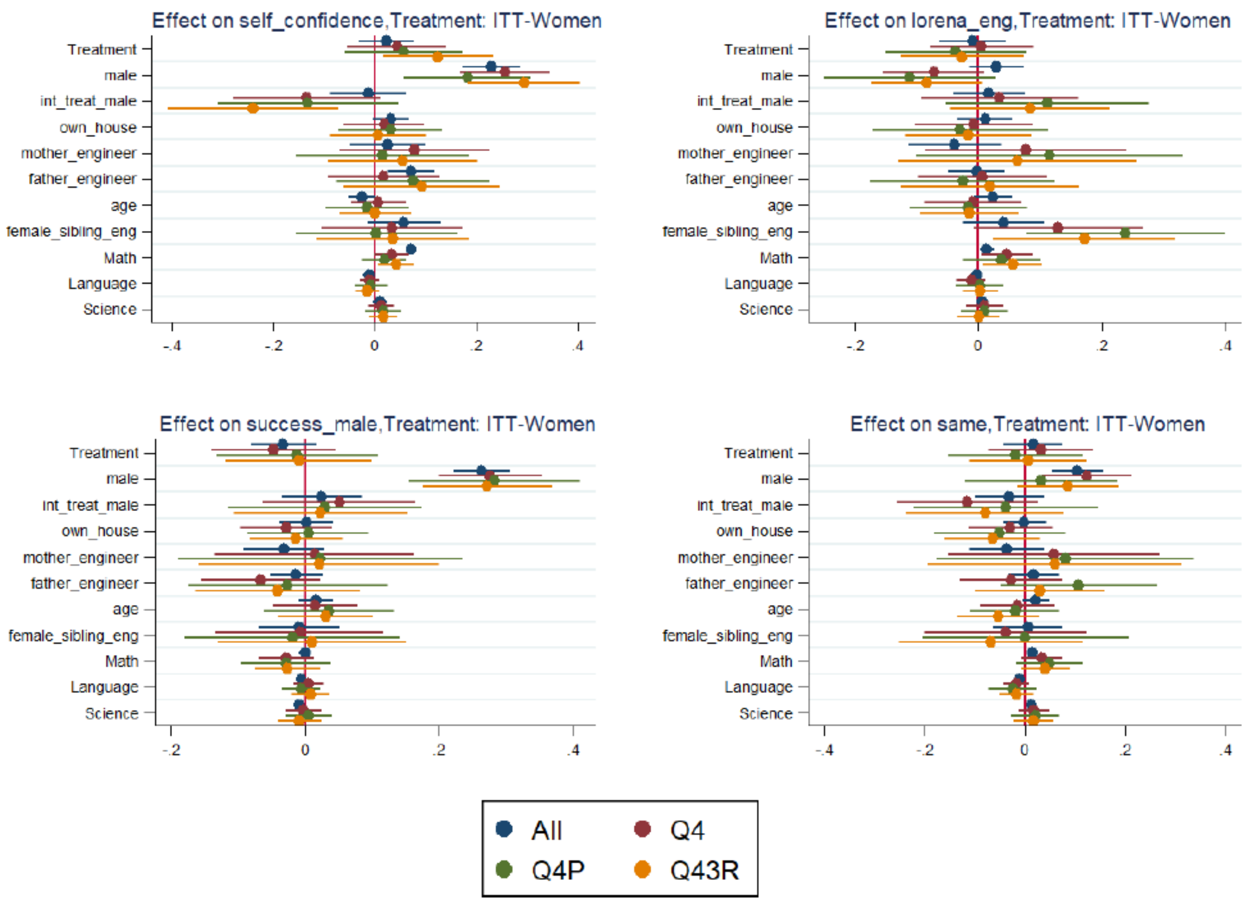
Notes: The figure shows the distribution of the math component of the ECE standardized examination of 2015 for both treatment and control schools.

Figure A7: Senior-year high school students- preference for fields of study



Notes: The figure shows the intent to treat (ITT) estimates for girls ("*Treatment*") and the effect of other covariates on senior-year students' preferences for fields of study: i) all types of Engineering, ii) the role models engineering majors (Industrial and Systems Engineering, Civil Engineering, and Mechanical/Electrical Engineering), iii) Non-STEM fields. Estimates for four subgroups are reported. *All* denotes the group for the entire sample of students, *Q4* includes only students in the upper quartile of baseline math scores, *Q4P* includes students in the upper quartile of baseline math scores and attending schools in Piura, *Q43R* incorporates students in the upper quartile of baseline math scores and attending schools in Piura, Tumbes, and Lambayeque (3 regions). Horizontal spikes denote 95% confidence intervals.

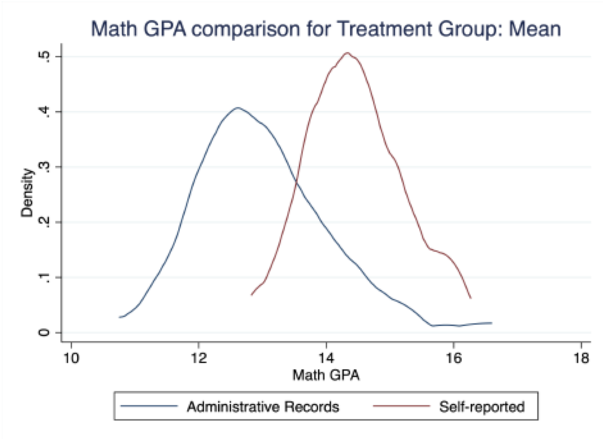
Figure A8: Senior-year high school students- perceptions



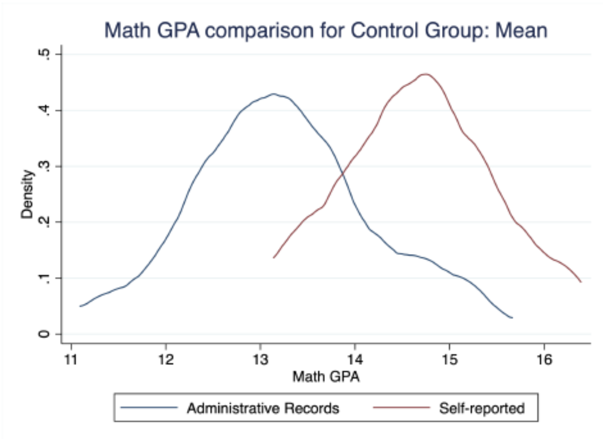
Notes: The figure shows the intent to treat (ITT) estimates for girls ("*Treatment*") and the effect of other covariates on senior-year students' perceptions: i) self-confidence in having aptitude and skills to pursue an engineering major, ii) recommending engineering to a hypothetical female friend (Lorena), iii) attributing success to men in engineering fields, iv) suggesting the same career to a hypothetical male and a hypothetical female friend. Estimates for four subgroups are reported. *All* denotes the group for the entire sample of students, *Q4* includes only students in the upper quartile of baseline math scores, *Q4P* includes students in the upper quartile of baseline math scores and attending schools in Piura, *Q43R* incorporates students in the upper quartile of baseline math scores and attending schools in Piura, Tumbes, and Lambayeque (3 regions). Horizontal spikes denote 95% confidence intervals.

Figure A9: Over-reporting by treatment status

(a) 10th Grade Math GPA- Treatment Group



(b) 10th Grade Math GPA- Control Group



Notes: The figure shows the over-reporting of math scores using i) follow-up survey and ii) 'Sistema de Información de Apoyo a la Gestión de la Institución Educativa (SIAGIE)' administrative records from the Ministry of Education (MINEDU). Panel (a) and (b) show the distribution of 10th grade school math scores for the treatment and the control group, respectively.

Table A1: Effect on students' preference for engineering: including covariates

Sample:	(1) Full	(2) 4Q	(3) AM	(4) BM	(5) 1Q	(6) 4Q3R
Treatment	0.0179	-0.00162	-0.0170	0.0416	0.0196	-0.0430
	(0.0237)	(0.0489)	(0.0367)	(0.0285)	(0.0355)	(0.0501)
Interaction	-0.00797	0.0928	0.0628	-0.0504	-0.0387	0.174***
(Treatment*female)	(0.0269)	(0.0593)	(0.0384)	(0.0328)	(0.0398)	(0.0651)
Female	-0.261***	-0.307***	-0.321***	-0.226***	-0.195***	-0.354***
gender (female=1)	(0.0184)	(0.0385)	(0.0310)	(0.0224)	(0.0251)	(0.0422)
own_house	0.0311*	0.0734	0.0571*	0.0168	-0.00729	0.0709
	(0.0164)	(0.0482)	(0.0328)	(0.0169)	(0.0260)	(0.0581)
mother_engineer	0.0364	0.0407	0.0384	0.0277	0.0227	0.0372
	(0.0290)	(0.0987)	(0.0555)	(0.0334)	(0.0456)	(0.126)
father_engineer	0.0412**	0.0876*	0.0609*	0.0239	0.00108	0.0941
	(0.0202)	(0.0442)	(0.0336)	(0.0238)	(0.0284)	(0.0623)
age	-0.0357***	-0.000528	-0.0200	-0.0422**	-0.0139	-0.00987
(in years)	(0.0124)	(0.0398)	(0.0235)	(0.0175)	(0.0196)	(0.0465)
has a female sibling engineer	0.0635**	0.167***	0.0728	0.0538	0.0370	0.218***
	(0.0246)	(0.0601)	(0.0460)	(0.0329)	(0.0423)	(0.0586)
Math	0.0510***	0.0290	0.0372**	0.0490***	0.0277***	0.0287
(10th grade math GPA)	(0.00383)	(0.0207)	(0.0147)	(0.00536)	(0.00922)	(0.0229)
Language	-0.0235***	-0.0328***	-0.0294***	-0.0206***	-0.0215***	-0.0200
(10th grade spanish GPA)	(0.00638)	(0.0120)	(0.0110)	(0.00504)	(0.00593)	(0.0143)
Science	-0.00507	-0.00310	-0.0105	-0.00350	-0.00640	-0.00208
(10th grade science GPA)	(0.00554)	(0.0114)	(0.00999)	(0.00567)	(0.00569)	(0.0122)
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Parent education FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,580	939	1,585	2,995	1,437	691
Adjusted R^2	0.161	0.136	0.144	0.143	0.118	0.141

Notes: This table reports the ITT estimates of the role model interventions on grade 11 students' preferences for engineering, including the estimates on covariates. The regression controls for city fixed effects and parental education fixed effects. Standard errors are clustered at the unit of randomization (school). 4Q corresponds to the sample of students in the top 25 percentile of baseline math scores, AM for students above the 50 percentile, BM for students below median or at the 50 percentile, 1Q for students in the bottom 25 percentile, and 4Q3R includes students in the upper quartile, and who are attending schools in three main regions (Piura, Tumbes, and Lambayeque). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A2: The effect of exposure to role models on students' preference for engineering by quartile of math performance: LATE

Outcome: Prefer Engineering	Control group mean female (1)	Treatment effect (LATE) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (LATE) male (5)	Standard error (6)	N (7)	Diff (LATE) p-value (8)
Panel A: Full Sample								
Q1	0.076	-0.020	0.017	0.271	0.021	0.038	1437	0.333
Q2	0.138	0.001	0.027	0.403	0.071	0.044	1558	0.174
Q3	0.194	-0.002	0.042	0.546	-0.073	0.070	646	0.347
Q4	0.205	0.097**	0.045	0.554	-0.002	0.052	939	0.121
Panel B: Main Regions								
Q1	0.068	-0.012	0.017	0.251	0.062	0.038	1132	0.093
Q2	0.138	0.012	0.029	0.389	0.068	0.051	1246	0.329
Q3	0.195	0.012	0.045	0.527	-0.008	0.076	515	0.821
Q4	0.175	0.139***	0.046	0.573	-0.046	0.053	691	0.010

Notes: This table reports the local average treatment effects (LATE) estimates for girls and the LATE for boys on preferences for engineering, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the LATE for females and males, respectively. The estimates are obtained from a two-stage least squares (2SLS) using treatment assignment as an instrument for treatment receipt. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city and it includes covariates. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table A3: Robustness check: high-ability math and science

Outcome: Prefer Engineering	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
top 25 M & S	0.184	0.090	0.069	0.506	0.005	0.092	395	0.443
top 25 M not S	0.225	0.083	0.053	0.581	-0.034	0.061	544	0.148
top 25 S not M	0.173	-0.161**	0.070	0.15	0.051	0.111	189	0.122
Panel B: Main Regions								
top 25 M & S	0.129	0.214***	0.074	0.582	-0.052	0.099	286	0.025
top 25 M not S	0.220	0.071	0.057	0.574	-0.085	0.070	405	0.091
top 25 S not M	0.211	-0.171*	0.089	0.192	0.033	0.143	135	0.278

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on preferences for engineering, separately by different groups of students based on skills in math (M) and science (S). Estimates correspond to i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A4: Robustness check: average school ECE math scores

Outcome: Prefer Engineering	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.076	-0.018	0.017	0.271	0.019	0.035	1434	0.356
Q2	0.138	0.018	0.026	0.403	0.078*	0.044	1539	0.219
Q3	0.194	0.013	0.042	0.546	-0.060	0.068	624	0.303
Q4	0.205	0.097**	0.041	0.554	0.018	0.047	907	0.174
Panel B: Main Regions								
Q1	0.068	-0.014	0.017	0.251	0.057	0.035	1129	0.093
Q2	0.138	0.036	0.028	0.389	0.078	0.051	1227	0.451
Q3	0.195	0.022	0.045	0.527	-0.009	0.076	493	0.711
Q4	0.175	0.132***	0.042	0.573	-0.029	0.052	659	0.016

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on preferences for engineering, separately for different subgroups of students based on self-reported baseline math scores. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) and it controls for average school 2015 ECE math scores. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A5: The effect on students' preference for engineering: School-UDEP distance, women in the top 25 percentile

VARIABLES	(1) Eng	(2) Eng	(3) Eng	(4) Eng	(5) Eng	(6) Eng
uddistreat	-0.142*	-0.161**	-0.156**	-0.162**	-0.167**	-0.167**
(Treatment*distanceAMUDEP)	(0.0747)	(0.0748)	(0.0745)	(0.0774)	(0.0802)	(0.0810)
distanceAMUDEP	0.119	0.109	0.109	0.192	0.181	0.181
	(0.101)	(0.115)	(0.114)	(0.120)	(0.125)	(0.125)
Treatment	0.148***	0.168***	0.162***	0.170***	0.172***	0.172***
ITT near schools	(0.0491)	(0.0481)	(0.0484)	(0.0527)	(0.0568)	(0.0570)
Treatment + uddistreat	0.006	0.006	0.006	0.008	0.005	0.006
ITT far schools						
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Parent Engineer	No	Yes	Yes	Yes	Yes	Yes
Own house	No	No	Yes	Yes	Yes	Yes
Parent Education FE	No	No	No	Yes	Yes	Yes
Baseline scores in 10th grade	No	No	No	No	Yes	Yes
Student's age (in years)	No	No	No	No	No	Yes
Has female sibling engineer	No	No	No	No	No	Yes
Observations	553	525	524	519	517	516
Adjusted R^2	0.034	0.038	0.041	0.043	0.044	0.045

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to female high ability students (fourth quartile of baseline math scores), who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. Column 1 reports the ITT estimates without covariates. Covariates are included from column 2 to column 6. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A6: The effect on students' preference for engineering: school-UDEP distance, men in the bottom 25 percentile

VARIABLES	(1) Eng	(2) Eng	(3) Eng	(4) Eng	(5) Eng	(6) Eng
uddistreat	-0.133**	-0.120	-0.117	-0.132*	-0.138**	-0.128*
(Treatment*distanceAMUDEP)	(0.0649)	(0.0733)	(0.0742)	(0.0714)	(0.0674)	(0.0694)
distanceAMUDEP	-0.0239	-0.0864	-0.0882	-0.0202	0.0331	0.133
	(0.0978)	(0.0873)	(0.0878)	(0.0850)	(0.0811)	(0.0857)
Treatment	0.111**	0.101*	0.101*	0.114**	0.0872*	0.0793*
ITT near schools	(0.0463)	(0.0521)	(0.0519)	(0.0481)	(0.0473)	(0.0472)
Treatment + uddistreat						
ITT far schools	-0.022	-0.019	-0.015	-0.018	-0.050	-0.049
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Parent Engineer	No	Yes	Yes	Yes	Yes	Yes
Own house	No	No	Yes	Yes	Yes	Yes
Parent Education FE	No	No	No	Yes	Yes	Yes
Baseline scores in 10th grade	No	No	No	No	Yes	Yes
Student's age (in years)	No	No	No	No	No	Yes
Has female sibling engineer	No	No	No	No	No	Yes
Observations	703	654	652	639	637	627

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to men in the bottom quartile of baseline math scores, who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. Column 1 reports the ITT estimates without covariates. Covariates are included from column 2 to column 6. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A7: The effect on students' preference for engineering: school-UDEP distance, men

VARIABLES	(1) Eng	(2) Eng	(3) Eng	(4) Eng	(5) Eng	(6) Eng
uddistreat	-0.111**	-0.115**	-0.111**	-0.101**	-0.109**	-0.101**
(Treatment*distanceAMUDEP)	(0.0466)	(0.0496)	(0.0480)	(0.0462)	(0.0497)	(0.0487)
distanceAMUDEP	-0.138**	-0.148**	-0.146**	-0.105	-0.0923	-0.0666
	(0.0630)	(0.0715)	(0.0715)	(0.0745)	(0.0612)	(0.0624)
Treatment	0.105***	0.104***	0.100***	0.0931***	0.0773**	0.0759**
ITT near schools	(0.0345)	(0.0355)	(0.0339)	(0.0310)	(0.0334)	(0.0323)
Treatment + uddistreat	-0.007	-0.011	-0.011	-0.007	-0.032	-0.025
ITT far schools						
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Parent Engineer	No	Yes	Yes	Yes	Yes	Yes
Own house	No	No	Yes	Yes	Yes	Yes
Parent Education FE	No	No	No	Yes	Yes	Yes
Baseline scores in 10th grade	No	No	No	No	Yes	Yes
Student's age (in years)	No	No	No	No	No	Yes
Has female sibling engineer	No	No	No	No	No	Yes
Observations	2,238	2,116	2,108	2,070	2,023	1,994
Adjusted R^2	0.012	0.012	0.013	0.020	0.080	0.081

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to men, who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. Column 1 reports the ITT estimates without covariates. Covariates are included from column 2 to column 6. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**} p < 0.05, ^{*} p < 0.1$.

Table A8: The effect on students' preference for engineering: school-UDEP distance, women

VARIABLES	(1) Eng	(2) Eng	(3) Eng	(4) Eng	(5) Eng	(6) Eng
uddistreat	-0.0757***	-0.0734**	-0.0744***	-0.0759***	-0.0516*	-0.0419
(Treatment*distanceAMUDEP)	(0.0274)	(0.0280)	(0.0279)	(0.0278)	(0.0307)	(0.0307)
distanceAMUDEP	-0.0145	-0.0369	-0.0426	-0.00704	-0.000503	-0.00206
	(0.0491)	(0.0509)	(0.0499)	(0.0531)	(0.0590)	(0.0556)
Treatment	0.0406**	0.0385**	0.0390**	0.0394**	0.0297*	0.0266
ITT near schools	(0.0172)	(0.0176)	(0.0172)	(0.0178)	(0.0176)	(0.0175)
Treatment + uddistreat	-0.035	-0.035	-0.035	-0.037*	-0.022	-0.015
ITT far schools						
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes	Yes
Parent Engineer	No	Yes	Yes	Yes	Yes	Yes
Own house	No	No	Yes	Yes	Yes	Yes
Parent Education FE	No	No	No	Yes	Yes	Yes
Baseline scores in 10th grade	No	No	No	No	Yes	Yes
Student's age (in years)	No	No	No	No	No	Yes
Has female sibling engineer	No	No	No	No	No	Yes
Observations	2,918	2,756	2,748	2,713	2,616	2,586
Adjusted R^2	0.009	0.008	0.008	0.014	0.051	0.052

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to women, who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. Column 1 reports the ITT estimates without covariates. Covariates are included from column 2 to column 6. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table A9: The effect on students' preference for engineering: school-UDEP distance, women (continued)

VARIABLES	(1) 4Q3R	(2) AM	(3) AM3R	(4) BM	(5) BM3R
uddistreat	-0.132	-0.0965*	-0.0785	-0.0146	0.0171
(Treatment*distanceAMUDEP)	(0.0862)	(0.0520)	(0.0557)	(0.0297)	(0.0416)
distanceAMUDEP	-0.920***	0.247**	-0.510***	-0.0720	-0.182***
	(0.0922)	(0.0958)	(0.0411)	(0.0634)	(0.0157)
Treatment	0.173***	0.0904***	0.0925***	-0.00343	-0.00375
ITT near schools	(0.0545)	(0.0324)	(0.0325)	(0.0136)	(0.0141)
Treatment + uddistreat	0.041	-0.006	0.014	-0.018	0.013
ITT far schools					
Observations	387	896	701	1,690	1,390
Adjusted R^2	0.051	0.039	0.041	0.029	0.038

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to female students, who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. "4Q3R" stands for upper quartile of baseline math scores in three main regions, "AM" stands for baseline math scores above the median, "AM3R" includes individuals with baseline math scores above the median and in three main regions, "BM" stands for baseline math scores below median, and "BM3R" denotes below median of baseline math scores in three main regions. Covariates include baseline scores in 10th grade, student's age in years, mother or father engineer indicator, ownership of house, parental education fixed effects, and an indicator for having a sibling engineer. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$.

Table A10: The effect on students' preference for engineering: school-UDEP distance, men (continued)

VARIABLES	(1) 1Q3R	(2) AM	(3) AM3R	(4) BM	(5) BM3R
uddistreat	-0.0479	-0.0949	-0.131	-0.0894	-0.113*
(Treatment*distanceAMUDEP)	(0.0664)	(0.0713)	(0.0933)	(0.0591)	(0.0659)
distanceAMUDEP	0.430***	-0.149**	-0.165***	0.0333	-0.0648**
	(0.0985)	(0.0627)	(0.0427)	(0.0998)	(0.0253)
Treatment	0.0784	0.0390	0.0383	0.0880**	0.0880**
ITT near schools	(0.0474)	(0.0469)	(0.0487)	(0.0439)	(0.0438)
Treatment + uddistreat	0.030	-0.056	-0.092	-0.001	-0.025
ITT far schools					
Observations	473	689	505	1,305	988
Adjusted R^2	0.046	0.011	0.007	0.075	0.073

Notes: This table reports the intent to treat (ITT) estimates on students' career preferences for engineering. The sample is restricted to male students, who answered the survey. uddistreat is the interaction term between our treatment variable and a dummy variable "distanceAMUDEP" that equals one if the school distance from UDEP is above the sample median, and zero otherwise. "1Q3R" stands for first quartile of baseline math scores in three main regions, "AM" stands for baseline math scores above the median, "AM3R" includes individuals with baseline math scores above the median and in three main regions, "BM" stands for baseline math scores below median, and "BM3R" denotes below median of baseline math scores in three main regions. Covariates include baseline scores in 10th grade, student's age in years, mother or father engineer indicator, ownership of house, parental education fixed effects, and an indicator for having a sibling engineer. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A11: Heterogeneous effects by type of engineering: only girls

	Preference for Engineering: Girls					
	Treatment effect (ITT)					
	Full Sample (1)	Above median (2)	4th Quartile (3)	Above median 3R (4)	4th Quartile 3R (5)	4th Quartile No 3R (6)
Industrial Engineering and Systems	0.015	0.032*	0.021	0.049***	0.040	-0.032
	(0.010)	(0.018)	(0.026)	(0.018)	(0.025)	(0.061)
Civil Engineering	0.009	0.026*	0.052***	0.022	0.063***	0.021
	(0.007)	(0.015)	(0.019)	(0.018)	(0.021)	(0.038)
Electrical and Mechanical Engineering	0.003	0.008	0.012	0.008	0.009	0.019
	(0.003)	(0.008)	(0.013)	(0.010)	(0.017)	(0.015)
N	2918	974	553	757	414	139

Notes: This table reports the treatment effects estimates on girls' preferences for Engineering by major, for different groups of students.The data is from a post-visit survey. Students' academic performance in math is measured by the students' score on math the previous year corresponding to grade 10. Intent-to-Treat estimates are displayed. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A12: Balance Test: Inflation of Math GPA

	Control Group (1)	Treatment Group (2)	Difference T-C (3)	p-value (4)
Over-reporting math grades	1.402	1.433	0.031	0.784
Over-reporting math grades for girls	1.225	1.285	0.060	0.626
Over-reporting math grades for boys	1.587	1.558	-0.029	0.845

Notes: This table shows the average school inflation of 10th math GPA in our follow-up sample. Column 1 and column 2 report the sample mean in the control and treatment group, respectively. Column 3 displays the estimate on the treatment dummy in a regression of each variable (Over-reporting math grades) on treatment. P-values for the statistically significance of the estimate are shown in column (4). Information comes from a follow-up survey implemented in 18 cities of Peru to senior-year high school students in November 2018 and from administrative records of 'Sistema de Información de Apoyo a la Gestión de la Institución Educativa' (SIAGIE) 2017- MINEDU.

Table A13: Balance Test: School Attendance rates

	Control Group (1)	Treatment Group (2)	Difference T-C (3)	p-value (4)
School attendance rates in 2018	0.866	0.950	0.071	0.389
School attendance rates in 2018 for girls	0.874	1.019	0.115	0.384
School attendance rates in 2018 for boys	0.866	0.926	0.054	0.414

Notes: This table shows the average school attendance rates in our follow-up sample. Column 1 and column 2 report the sample mean in the control and treatment group, respectively. Column 3 displays the estimate on the treatment dummy in a regression of each variable (school attendance) on treatment. We include city fixed effects. P-values for the statistically significance of the estimate are shown in column (4). Information comes from a follow-up survey implemented in 18 cities of Peru to senior-year high school students in November 2018 and from administrative records of 'Sistema de Información de Apoyo a la Gestión de la Institución Educativa' (SIAGIE) 2018- MINEDU.

Table A14: Balance Test: Non-reporting answers to questions in the follow-up survey

	Control Group (1)	Treatment Group (2)	Difference T-C (3)	p-value (4)
Missing information in 2018	5.981	8.298	2.210	0.154
Missing information in 2018 for girls	2.924	3.609	0.634	0.452
Missing information in 2018 for boys	3.511	5.238	1.791	0.085

Notes: This table shows the non-reporting behavior observed in the follow-up survey between treatment and control schools. The variable representing missing information is the school's number of students who did not answer at least one question related to career preferences and perceptions (i.e. preferences for engineering, self-confidence, males' success in the profession, knowledge about engineering types, gender stereotypes, among others). Column 1 and column 2 report the sample mean in the control and treatment group, respectively. Column 3 displays the estimate on the treatment dummy in a regression of each variable on treatment. We include city fixed effects. P-values for the statistically significance of the estimate are shown in column (4). Information comes from a follow-up survey implemented in 18 cities of Peru to senior-year high school students in November 2018.

Table A15: Multiple Hypothesis Test- Anderson: The effect of exposure to role models on students' preference for engineering and perceptions for high ability girls in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Pref-Eng (1)	Self-Conf (2)	Male Success (3)	Lorena-Eng (4)	Count-Eng (5)
ITT female: Treatment + Interaction	0.131***	0.125**	-0.009	-0.026	-0.203**
p-values	0.0039	0.0248	0.8658	0.6005	0.0100
q-values	0.014	0.040	0.909	0.801	0.020
Number of observations (N)	691	708	674	697	710
Dep. Variable:	Salary (6)	Eng Role Model (7)	NoSTEM (8)		
ITT female: Treatment + Interaction	0.007	0.134***	-0.134***		
p-values	0.9084	0.0018	0.0052		
q-values	0.909	0.014	0.014		
Number of observations (N)	707	691	691		

Notes: This table reports the intent to treat (ITT) estimates on female students' career preferences for engineering and perceptions. The sample is restricted to high ability girls (fourth quartile of baseline math scores) in schools located in Piura/Lambayeque/Tumbes, who answered the survey. The regression controls for city fixed effects and covariates. p-values and q-values are reported. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A16: Multiple Hypothesis Test- Anderson: The effect of exposure to role models on students' preference for engineering and perceptions for low ability boys in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Pref-Eng (1)	Self-Conf (2)	Male Success (3)	Lorena-Eng (4)	Count-Eng (5)
ITT male: Treatment + Interaction	0.059	0.034	-0.026	-0.039	-0.142
p-values	0.107	0.505	0.471	0.436	0.137
q-values	0.362	0.578	0.578	0.578	0.362
Number of observations (N)	1132	1190	1126	1172	1198
Dep. Variable:	Salary (6)	Eng Role Model (7)	NoSTEM (8)		
ITT male: Treatment + Interaction	0.022	0.051	-0.066*		
p-values	0.664	0.181	0.094		
q-values	0.664	0.362	0.362		
Number of observations (N)	1187	1132	1132		

Notes: This table reports the intent to treat (ITT) estimates on male students' career preferences for engineering and perceptions. The sample is restricted to low ability boys (first quartile of baseline math scores) in schools located in Piura/Lambayeque/Tumbes, who answered the survey. The regression controls for city fixed effects and covariates. p-values and q-values are reported. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A17: The effect of exposure to role models on students' preference for engineering by quartile of math performance

Outcome: Prefer Engineering	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.076	-0.019	0.016	0.271	0.020	0.035	1437	0.333
Q2	0.138	0.001	0.025	0.403	0.067	0.042	1558	0.174
Q3	0.194	-0.002	0.039	0.546	-0.069	0.066	646	0.347
Q4	0.205	0.091**	0.042	0.554	-0.002	0.049	939	0.121
Panel B: Main Regions								
Q1	0.068	-0.012	0.016	0.251	0.059	0.036	1132	0.093
Q2	0.138	0.011	0.027	0.389	0.064	0.048	1246	0.329
Q3	0.195	0.011	0.042	0.527	-0.008	0.072	515	0.821
Q4	0.175	0.131***	0.044	0.573	-0.043	0.050	691	0.010

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on preferences for engineering, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A18: The effect of exposure to role models on students' perceptions by quartile of math performance: self-confidence

Outcome: Self-confidence	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.197	0.002	0.035	0.391	0.045	0.042	1509	0.383
Q2	0.346	0.022	0.038	0.564	0.045	0.041	1612	0.635
Q3	0.488	0.045	0.052	0.773	-0.062	0.069	658	0.205
Q4	0.580	0.044	0.050	0.835	-0.090**	0.044	960	0.070
Panel B: Main Regions								
Q1	0.194	0.003	0.040	0.395	0.034	0.050	1190	0.588
Q2	0.343	0.034	0.041	0.565	0.055	0.046	1290	0.710
Q3	0.515	0.047	0.058	0.740	0.019	0.084	522	0.788
Q4	0.551	0.125**	0.054	0.853	-0.116**	0.053	708	0.006

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on students' self-confidence in their aptitude and skills to pursue an engineering major, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A19: The effect of exposure to role models on students' perceptions by quartile of math performance: success exclusively for men in the sector

Outcome: Male Successful	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.633	-0.037	0.040	0.391	-0.040	0.032	1400	0.951
Q2	0.605	-0.050	0.039	0.564	0.011	0.029	1524	0.164
Q3	0.629	0.017	0.050	0.773	-0.004	0.050	624	0.774
Q4	0.557	-0.047	0.047	0.835	0.004	0.041	894	0.375
Panel B: Main Regions								
Q1	0.651	-0.061	0.044	0.395	-0.026	0.036	1126	0.565
Q2	0.609	-0.045	0.043	0.565	-0.011	0.032	1233	0.473
Q3	0.648	-0.007	0.056	0.740	-0.007	0.063	499	0.999
Q4	0.567	-0.009	0.055	0.853	0.014	0.048	674	0.723

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on students' perceptions of males successfulness in engineering, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table A20: The effect of exposure to role models on students' perceptions by quartile of math performance: gender stereotypes

Outcome: Engineering to Lorena	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.466	-0.046	0.042	0.516	-0.046	0.042	1473	0.997
Q2	0.457	-0.001	0.042	0.508	0.054	0.037	1589	0.354
Q3	0.535	0.048	0.057	0.579	-0.019	0.074	649	0.427
Q4	0.579	0.005	0.042	0.527	0.040	0.051	941	0.588
Panel B: Main Regions								
Q1	0.464	-0.035	0.049	0.492	-0.039	0.050	1172	0.940
Q2	0.448	0.003	0.047	0.479	0.086*	0.044	1270	0.235
Q3	0.515	0.039	0.066	0.552	-0.006	0.092	520	0.644
Q4	0.589	-0.026	0.050	0.520	0.057	0.049	697	0.205

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on students' recommending engineering to Lorena (hypothetical female friend), separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table A21: The effect of exposure to role models on students' perceptions by quartile of math performance: knowledge of engineering fields

Outcome: Types of engineering listed	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	4.348	-0.075	0.085	4.177	-0.097	0.088	1518	0.847
Q2	4.334	-0.055	0.080	4.330	0.026	0.069	1621	0.396
Q3	4.502	0.032	0.089	4.484	0.085	0.119	661	0.698
Q4	4.610	-0.096	0.070	4.526	0.040	0.088	963	0.163
Panel B: Main Regions								
Q1	4.324	-0.035	0.098	4.205	-0.142	0.094	1198	0.379
Q2	4.334	-0.004	0.085	4.346	0.030	0.077	1296	0.763
Q3	4.515	0.018	0.092	4.458	0.157	0.146	525	0.379
Q4	4.654	-0.203**	0.077	4.565	0.005	0.083	710	0.063

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on students' number of engineering fields listed, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A22: The effect of exposure to role models on students' perceptions by quartile of math performance: earnings expectations

Outcome: Salary (in logarithm)	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	8.196	-0.056	0.041	8.168	-0.003	0.044	1499	0.353
Q2	8.194	-0.032	0.037	8.182	0.051	0.037	1613	0.072
Q3	8.154	-0.007	0.043	8.239	-0.087	0.054	655	0.201
Q4	8.235	-0.007	0.052	8.230	-0.056	0.043	953	0.465
Panel B: Main Regions								
Q1	8.199	-0.064	0.045	8.168	0.022	0.050	1187	0.201
Q2	8.188	0.002	0.038	8.176	0.095**	0.041	1291	0.061
Q3	8.168	-0.008	0.047	8.244	-0.090	0.068	523	0.270
Q4	8.226	0.007	0.062	8.266	-0.078	0.052	707	0.249

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on students' knowledge about earnings associated with engineering careers, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A23: Students' preference for the role models' majors by quartile of math performance

Outcome: Any three types of engineering	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.042	-0.006	0.015	0.220	0.013	0.036	1437	0.635
Q2	0.101	0.012	0.023	0.321	0.046	0.038	1558	0.443
Q3	0.146	0.015	0.032	0.496	-0.064	0.061	646	0.225
Q4	0.142	0.104**	0.040	0.512	-0.030	0.048	939	0.019
Panel B: Main Regions								
Q1	0.034	0.000	0.016	0.202	0.051	0.038	1132	0.224
Q2	0.104	0.023	0.025	0.306	0.052	0.042	1246	0.559
Q3	0.152	0.018	0.034	0.484	-0.011	0.073	515	0.725
Q4	0.132	0.134***	0.041	0.534	-0.074	0.048	691	0.001

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on preferences for the role models' majors offered at UDEP (industrial engineering, civil engineering, or mechanical engineering), separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A24: The effect of exposure to role models on students' preference for non-stem fields by quartile of math performance

Outcome: Prefer Non-STEM	Control group mean female (1)	Treatment effect (ITT) female (2)	Standard error (3)	Control group mean male (4)	Treatment effect (ITT) male (5)	Standard error (6)	N (7)	Diff (ITT) p-value (8)
Panel A: Full Sample								
Q1	0.911	0.020	0.018	0.723	-0.032	0.037	1437	0.222
Q2	0.850	0.001	0.027	0.583	-0.065	0.043	1558	0.187
Q3	0.806	-0.009	0.040	0.445	0.063	0.065	646	0.319
Q4	0.795	-0.097**	0.045	0.430	-0.013	0.043	939	0.135
Panel B: Main Regions								
Q1	0.915	0.017	0.018	0.741	-0.066*	0.039	1132	0.073
Q2	0.847	-0.005	0.029	0.591	-0.051	0.049	1246	0.397
Q3	0.805	-0.023	0.043	0.462	0.000	0.070	515	0.785
Q4	0.825	-0.134***	0.046	0.416	0.020	0.045	691	0.017

Notes: This table reports the intent to treat (ITT) estimates for girls and the ITT for boys on preferences for non-STEM fields, separately by quartile of performance in math and for i) Full sample of schools (Panel A), ii) only schools located in main regions (Piura, Tumbes, Lambayeque), Panel B. Column 1 and column 4 show the average value for female and male students in the control group, respectively. Column 2 and column 5 report the intent to treat estimates (ITT) for females and males, respectively. The estimates are obtained from a regression following equation (1) including covariates. p-value for the difference in means test among males and females is reported in column 8. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are reported in column 3 and 6. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Table A25: The effect of exposure to role models on students' preference for non-STEM fields in Piura/Lambayeque/Tumbes schools

Dep. Variable:	Prefer Non-STEM			
Sample:	Q1 math (1)	Q2 math (2)	Q3 math (3)	Q4 math (4)
Treatment	-0.066*	-0.051	0.000	0.020
	(0.039)	(0.049)	(0.070)	(0.045)
Female	0.179***	0.250***	0.377***	0.371***
	(0.029)	(0.040)	(0.061)	(0.040)
Interaction	0.083*	0.046	-0.023	-0.153**
(Treatment*female)	(0.046)	(0.054)	(0.086)	(0.063)
ITT female: Treatment + Interaction	0.017	-0.005	-0.023	-0.133***
City FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Number of observations (N)	1132	1246	515	691
Adjusted R^2	0.121	0.134	0.167	0.159
Mean Dv (Treatment==0)	0.91	0.85	0.80	0.82

Notes: This table reports the intent to treat (ITT) estimates on students' preferences for non-STEM fields, separately by quartile of performance in math. The sample is restricted to students in schools located in Piura/Lambayeque/Tumbes. Control variables include: has an engineering parent, owns house, parental education FE, baseline scores in 10th grade, age and having an engineer sibling. The regression controls for city fixed effects since the randomization was stratified by city. Standard errors clustered at the unit of randomization (school) are shown in parentheses. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

Survey Instruments

Student survey: survey about preferences and perceptions of fields of study among senior-year high school students in Peru

Q1. School:
Q2. City:
Q3. Sex: 1. ☐ male 2. ☐ female
Q4. Age (In years completed):
Q5. Final course grade on Math in grade 10:
Q6. Final course grade on Language in grade 10:
Q7. Final course grade on Science in grade 10:
☺ If you do not remember exact grades please write an approximation.
☺ Now, we are going to ask easy questions about your career preferences. Remember that there is no correct or incorrect answer. Please respond to the following questions honestly.
Q8. Would you like to study at university after graduating from high school?
(Important: select only one option. If you are still undecided, select the option that comes close to what you would like to do)
1. ☐ Yes ⟶ (Go to question Q9 and continue the survey if your choice was "Yes")
2. ☐ No ⟶ (Go to question Q10 and continue the survey if your choice was "No")

Q9. Please write the name of the career you would like to study the most in any university. (If you are in doubt between several careers that you like the same please write the name of one of them)
Q10. Have you already decided at which university to study? (Select the option that applies)
1. ☐ Yes ⟶ (Go to question Q11 and continue the survey if your choice was "Yes")
2. ☐ No ⟶ (Go to question Q12 and continue the survey if your choice was "No")

Q11. Please answer questions Q11a, Q11b, and Q11c:

Q11a. Write the name of the university where you have decided to study:
Q11b. Write the name of the career that you are going to study at this university:
Q11c. Are you already enrolled or have you reserved a place in this university? (Select one option only and go to question Q12. Continue the survey)
1. ☐ Yes 2. ☐ No
☺ Read carefully each of the following questions, and answer according to your own view. Remember that there is no correct or incorrect answer.
Q12. Imagine that Javier and Lorena are two of your best friends. Both of them have a final course grade in Math and in Science of 20. Javier and Lorena are not sure which career to study. Which field of study would you suggest to each of them?
Field of study that you suggest to Lorena:
Field of study that you suggest to Javier:
Q13. One person studied Informatics Engineering in the best university in Peru. After having worked for more than 10 years in companies such as Microsoft, Facebook, IBM, and Google, this person started his/her own business. His/her company is one of the top five leading engineering companies in the country. In your opinion: (Please select one option only)
1. ☐ Even though this person can be male or female, it is more probable that is male.
2. ☐ Even though this person can be male or female, it is more probable that is female.
Q14. One type of engineering is civil engineering. Please list five other types of engineering: (If you do not remember another five types of Engineering, list the ones you remember and leave the other blanks unfilled)
Q15. One person graduated from the Industrial Engineering program offered by a university in Peru two years ago. Currently, the person is working. How much do you think the person earns per month? (Select one option only)
1. ☐ Less than 1000 soles 3. ☐ Between 2000 and 3000 soles 5. ☐ Between 4000 and 5000 soles
2. ☐ Between 1000 and 2000 soles 4. ☐ Between 3000 and 4000 soles 6. ☐

Between 5000 and 6000 soles
7. ☐ Between 6000 and 7000 soles 8. ☐ Between 7000 and 8000 soles 9. ☐ Between 8000 and 9000 soles
10. ☐ More than 9000 soles
Q16. Do you think you have the capacities and qualities to study Engineering at university? (Select one option only)
1. ☐ Yes, I have them 2. ☐ No, I don't have them 3. ☐ I don't know
☺ Next, we are going to ask you some easy questions about your parents. Please respond the best you can to the following questions:
Q17. Age of your father/ attorney in years completed:
Q18. Is your father/attorney an engineer? (Select the option that applies): 1. ☐ Yes 2. ☐ No
Q19. Please select the level of education of your father/attorney:
1. ☐ Primary education completed 3. ☐ Technical education incomplete 5. ☐ University incomplete
2. ☐ Secondary education completed 4. ☐ Technical education completed 6. ☐ University completed
Q20. Does your father/ attorney work?: 1. ☐ Yes 2. ☐ No
Q21. Age of your mother in years completed:
Q22. Is your mother an engineer? (Select the option that applies): 1. ☐ Yes 2. ☐ No
Q23. Please select the level of education of your mother:
1. ☐ Primary education completed 3. ☐ Technical education incomplete 5. ☐ University incomplete
2. ☐ Secondary education completed 4. ☐ Technical education completed 6. ☐ University completed
Q24. Does your mother work?: 1. ☐ Yes 2. ☐ No
☺ Now we are going to ask questions about your siblings. For each question cross the cell that corresponds:
Q25. How many siblings do you have in total? $\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q26. How many brothers do you have in total? $\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q27. How many sisters do you have in total? $\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$

Q28. How many of your brothers are currently studying at university?
$\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q29. How many of your sisters are currently studying at university? $\boxed{0}$
$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q30. How many of your brothers are currently studying engineering?
$\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q31. How many of your sisters are currently studying engineering?
$\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q32. How many of your brothers are engineers?
$\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
Q33. How many of your sisters are engineers?
$\boxed{0}$ $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{\geq 6}$
☺ Now, we are going to ask some easy questions about the household. Please answer them the best you can:
Q34. Does your family live in an own or rented house?: 1. ☐ Own 2. ☐ Rented 3. ☐ Other (Specify):
Q35. Is there a car or truck in your home? :
1. ☐ Yes → How many? $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{\geq 5}$
2. ☐ No
Q36. Is there a motorcycle in your home? :
1. ☐ Yes → How many? $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{\geq 5}$
2. ☐ No
Q37. Is there a TV in your home? :
1. ☐ Yes → How many? $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{\geq 5}$
2. ☐ No
Q38. Is there a computer or laptop in your home? :
1. ☐ Yes → How many? $\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{\geq 5}$
2. ☐ No
Q39. Do you have internet access at home? : 1. ☐ Yes 2. ☐ No
Q40. Have you gone on vacation with your family to any place in Peru this 2018? : 1. ☐ Yes 2. ☐ No
Q41. Have you traveled abroad with your family this 2018? : 1. ☐ Yes 2. ☐ No
☺ Finally, tell us whether did you register to take the University of Piura's

PAE test in 2018, and to what career did you apply in the PAE:

Q42. Did you register to take the University of Piura's PAE test this year 2018? (Select the option that applies) :

1. ☐ Yes (If "Yes" go to question Q43)
2. ☐ No (If "No", this is the end of the survey, thank you!)

Q43. To what career did you apply in the PAE test? (Please state the career that you selected when you registered to take the PAE test) :

Thank you!