Empowering the future of Science: Experimental evidence on STEM skills and attitudes in Peru

Fabiola Alba Vivar *

August, 2024

PRELIMINARY DRAFT

Abstract

I study a program designed to reduce the gender gap in STEM fields by providing a bundled intervention targeting the abilities and aspirations of girls in primary school in a larga city in Latin America. The program provides weekly science workshops which include lectures with top scientists, after-school activities and mentorship. I find that the program has no detectable effects on girls' academic performance in school, but sizable effects on several other outcomes, including confidence in their science abilities, perceptions of non-STEM majors, and time use. I also document a trade-off between after-school study time and time spent on personal projects, suggesting a relative gain in productivity given the null effects on academic performance and negative effects on effort.

*I have benefited from meetings and data provided by MacTec Peru and the support of Johanna Johnson, Wendy Patino and Javier Navarro. In previous versions of this paper, I received the support of Maria Rosales Rueda and Micaela Sviatschi. I am also gratefully acknowledge all the helpful comments and suggestions from Alex Eble, Sarah Cohodes, Matteo Magnaricotte, Giannina Vaccaro, Astrid Pineda, Daniel Sparks, and seminar participants at PUCP, AEFP and Columbia's Economic and Education Colloquium. Special thanks to our research assistants Lucia Carrillo, Sarita Ore, Maria Fernanda Jaramillo who collaborated at different stages of this project. I am also thankful to Cindy Rojas for her outstanding research analysis.

* Wake Forest University, *albavif@wfu.edu,* Winston-Salem, NC.

1 Introduction

Over the past 30 years, the gender gap in educational attainment has been reduced in many developed countries and more recently in developing ones (Bailey and Dynarski, 2011; Rosenzweig and Zhang, 2013) providing a positive outlook for young girls worldwide. Nonetheless, the gender gap in science, technology, engineering, and mathematics (STEM) education and careers strongly persists (Ellison and Swanson, 2010; Cheryan et al., 2015; Gonzalez De San Roman and De La Rica, 2016; Eble and Hu, 2022; Jayachandran, 2015). For example, in poorer regions like Africa and Latin America, the gender gap in mathematics achievement tends to favor boys as early as the end of the primary school (UNESCO, 2017). Despite having evidence suggesting that female participation, achievement, and progression in STEM are driven by multiple and overlapping factors such as social norms and gender stereotypes, there is still a lack of evidence on what type of policies can mitigate these gaps. In this paper, I examine the effects of an after-school program that aims to increase primary school girls' STEM skills, aspirations, and attitudes in a large city in South America.

I study the case of an educational program that randomly chooses girls to participate in a one-year program consisting of a weekly workshop with top American and Peruvian scientists and university professors. This is a unique intervention as it targets girls in primary school, a critical period before gaps in STEM abilities widen.These workshops, while they do not follow the school year curricula or provide any remedial education, aim to promote girls' natural interest in science while exposing them to successful role models. Conditional on applying, the program selects participants using a lottery, which provides an ideal setting to causally estimate the impact of this program. I analyze the 2016-2019 program cohorts by matching applicants ($\sim$ 2,800 girls) to administrative school records (enrollment and grades by subject). Additionally, I conducted an endline survey to measure attitudes, beliefs and aspirations. The empirical strategy leverages the lottery assignment of applicants by comparing post outcomes between participants versus non-participants conditional on applying. Furthermore, I also analyze the sample of girls applying to the program and find a high positive selection compared to the average population of eligible girls in that age cohort. Using administrative school data, I find that girls who applied to the program attend better schools and using data from the 2019 application data, I find that applicants' parents have significantly more years of education compared with the eligible population. Nevertheless, I am able to corroborate the validity of the randomization. The baseline sociodemographic and educational characteristics (before the intervention) are balanced between selected (treated) and non-selected (control) applicants, which validates the empirical strategy.

The main results suggest that this program had no significant effects on academic achievement measured with school grades, even two years after the program. However, the endline results show that girls who participated in the program are more confident about their science grades at school and when contrasting their perceptions about their performance with real data, I find that they are more overconfident relative to the applicants who were not chosen for the program. I also do not find significant effects on their perceptions of natural ability and effort. In particular, the signs of the impacts on effort are negative and consistent with the effects on time use: girls in the treatment group report spending less time on studying and school homework but they spend more time on personal projects. This suggests that girls in the treatment group might be trading off school-related time with personal development activities, and therefore, becoming more productive given that there are no significant changes on grades.

Regarding expectations and aspirations about their future, I asked girls about their aspirations for education and occupation and aggregate their responses into an aspiration index. I do not find evidence that their participation in the program changed those measures of aspirations, but it is worth highlighting that these expectations were already high as reported in our baseline collected in 2019. In addition, I asked girls about their perceived happiness if they were to study a list of different STEM and non-STEM majors. The results suggests that girls in the treatment group seem to have more pessimistic perceptions about any major in college but the effects are stronger for non-STEM majors, especially Law, Education, and Journalism. They also reported being very pessimistic about Mathematics and Architecture majors, which are STEM-related majors that are not covered in the workshops of the program. When it comes to major choice, I do not find significant effects on STEM-related majors. This could potentially be explained since girls who applied to the program were already interested in STEM at baseline, leaving the margin to change their perceptions very low. We confirm this hypothesis with some questions regarding social norms, where I do not find effects on attitudes towards marriage and ladylike careers. Nonetheless, I find that girls who participated in the program seem to be less likely to follow family advice and to talk to their parents about their educational future, suggesting a higher degree of independence. Finally, it is worth noting that since the endline was implemented during the COVID-19 pandemic, I also explored girls' mental health status at that time. Even when the coefficients are not significant, I find signs that girls in the treatment group might be experiencing higher levels of stress and frustration than those who did not participate in the program.

This paper contributes to several works of literature. First, the program provides a unique STEM intervention as it targets girls during primary school at ages 8 to 11, while most similar interventions focus on older students. This is an important feature since interventions at earlier years could potentially have bigger impacts and this program is implemented at a critical period before gaps in STEM abilities widen up (UNESCO, 2017). The literature has provided strong positive evidence on STEM programs targeting secondary or post-secondary women (Cohodes et al. (2022); Del Carpio and Guadalupe (2021); Kitchen et al. (2018); Moss-Racusin et al. (2018)). The closest work is from Cohodes et al. (2022) who implemented a randomized control trail (RCT) for a STEM summer program for underrepresented high school students in the US. Their results show positive effects on college enrollment, persistence, graduation rates, and likelihood of earning a STEM degree. Unlike their work, we evaluate a full academic year after-school program for younger girls whose prospects of attending college and making a career choice are still far away in their lives. Instead, we focus on short-term outcomes and evaluate changes in perceptions and general STEM interests. In that sense, we also contribute to a growing body of evidence that suggests that early interventions to this demographic group can be very effective as attitudes and soft skills are being formed and are, therefore, malleable (Alan and Ertac, 2018; Bandiera et al., 2020;

Dhar et al., 2019, 2022; Heckman and Rubinstein, 2001). I provide evidence on the effects of an after-school intervention targeting girls before they initiate adolescence in a large urban setting in a developing country. The results suggest that the intervention was not effective at this very early stage in life. However, I am not able to corroborate longer-term outcomes such as college and major choice.

Second, I study a novel intervention based on role models. There is evidence that interventions involving educational or professional role models can improve children's outcomes Beaman et al. (2012); Tanguy et al. (2014); Breda et al. (2023); Dennehy and Dasgupta (2017); Kearney and Levine (2020); Porter and Serra (2020). In the same spirit as external role models, some other studies have shown positive effects of the role of female teachers on girls' academic achievement and enrollment in STEM majors (Eble and Hu, 2020; Carrell et al., 2010; Griffith, 2014; Griffith and Main, 2019, 2021) and enrollment in STEM. Moreover, Carlana (2019), using gender-science implicit association tests, finds that teachers with stronger implicit stereotypes negatively affect math achievements of their female students. Thus, teachers may also be affecting students' outcomes by exposing them to different gender stereotypes. I complement this literature by studying the role of being exposed to top scientists who are external to the classroom setting during a full academic year. Although we are not able to disentangle the differential effects, this program provides evidence on both top role models (i.e a university professor in the US) and local role models (i.e Peruvian girls majoring in science in a local university). What is more, in this setting, 50 percent of the scientists mentors are female which reflects the program efforts to have parity even when in reality it is hard to find top female scientists in Peru, similar to most developing countries. The closest work to this paper is Breda et al. (2023), which finds that a one-hour visit of a female role model with a background in science has an impact on students' choice of field of study after high school graduation, particularly for high-achieving females. They also find an increase in the probability that a female student would enroll in a male-dominated STEM track in college. In contrast with their work, I focus on young girls in primary school and study the effects of a longer program in a developing country where the gender gap is even wider.

Third, a large body of research studied how aspirations and beliefs are formed and their effect on human capital investments (Tanguy et al., 2014; Akerlof and Kranton, 2000; Benabou and Tirole, 2011; Lybbert and Wydick, 2018). For example, it has been documented that a lack of math self-efficacy drives women's dropout from STEM majors (Saltiel, 2022). I complement this work by showing how external role models can affect these outcomes for girls at a critical period for human capital formation. In particular, I study how exposure to top scientists can affect girls' aspirations and beliefs and how this may translate into investment in skills, and academic performance. The results suggest that this intervention helped young girls be more productive at school: keeping their high grades while attending a full-year program and later devoting time to personal projects.

The next section describes the program and the intervention I am studying. Section 3 presents the experimental design and the validation of the randomization. Section 4 presents the results using administrative data and an endline survey and Section 5 concludes with a discussion of the results.

2 Background and Data

2.1 Description of the Intervention

The program was implemented in Lima, Peru's capital and a city of approximately 12 million people. The Peruvian context, similar to other developing countries, shows a very rough scenario for women. For example, the wage gender gap is 30% and it remained the same for over a decade before the COVID-19 pandemic, which only increased it. Even for highly educated women in STEM (women with a BA degree), the wage gender gap is 23%. What is more, the adversities for women do not come exclusively from earnings: many barriers and difficulties are also cultural. Lima is considered one of the worst cities in the world for women, comparable to New Delhi and Kampala.[1] As documented by Sviatschi and

[1] Plan International has documented experiences in cities worldwide, where Lima came out as one of the worst cities in the world for women in terms of safety. See report here.

Trako (2024), Peru is a country that has experienced a huge increment in gender violence, where the number of domestic violence cases registered in local police departments increased substantially: from 29,759 in 2002 to more than 60,000 in 2016. At the same time, numerous political and religious groups have emerged claiming their opposition to the *'gender ideology'* and the implementation of a school curriculum that includes a gender equality agenda [2].

In this context, a Peruvian NGO founded in 2012 and based in an elite private college, created a program that aimed to increase women's participation in scientific fields. Pushing their agenda for gender equality in STEM, they provide a safe and friendly environment for girls to demonstrate their curiosity and creativity for science. After a few years of improving their pilot, they were able to reach out to multiple districts in the city. Starting in 2016, they randomly choose 40 girls for a one-year program consisting of weekly workshops with American and Peruvian top scientists. Also, after the workshops, participants work in small groups under the mentoring of a big-sister fellow[3], following the *Discovery Peer Learning* method. This method is quite different from traditional classes: participants are invited to work along with the lecturer (who acts more like a leader rather than a teacher). During these workshops, creativity and critical thinking are strongly encouraged with constant debate and questioning. Additionally, no rankings or test scores are used in the workshops. Since this is not a remedial program, they do not follow or include topics from the Ministry of Education's curricula either. Most of the topics covered in class are university-level topics that were simplified for younger students. For example, the workshops include topics such as optics, DNA extraction, or crystallization processes.

I interviewed the program staff, and they shared with us their difficulties when it comes to recruiting volunteer staff and fellows. Even when they have aimed to reach a more diverse audience, and to expand their outreach outside Lima, there are still several barriers to reaching girls at the bottom of the income distribution. Despite this, the program had

[2] #ConMisHijosNoTeMetas is an active social movement that has spread to other countries like Colombia, Chile, and Spain. Their main claim is that governments should not teach children about gender and sexual orientation since this choice belongs to the parents. Their flag is blue and pink, symbolizing the color of masculine and feminine genders.

[3] This fellow is typically a Peruvian college STEM female student who worked leading a group of 5 girls during the entire year.

an increasing demand over the years. Anecdotally, elite private schools in Lima have often requested the NGO to implement their program in their schools and to pay for it but they are not able to expand it due to the high cost of the implementation. The average cost of one scholarship is 2,000 USD per year, compared with the average government expenditure of approximately 1,000 USD per year in primary school.

The program runs every year following the timeline shown in Table 1. It starts during the Peruvian summer break (January-March) with the application season. They advertise mainly on social media in January and February and implement school visits in March when the school year starts. The online applications are open until the last day of March, and the lottery is usually run in a public ceremony. All workshop activities are held every Saturday ($\sim 4 hours$ per session), from April to November, except during the national holidays or the Peruvian winter break (mid-July). Parents of selected participants are asked to sign an agreement where they committed to attending every workshop and were aware that after 3 unjustified missing classes, the students are invited to retire from the program. On average, one girl per year dropped out of the program within the first weeks. If a girl dropped out during the first 3 weeks, she was replaced with another randomly assigned girl. This study focused on all girls who finished the program at the end of the academic year.

Table 1: Official Timeline of the Program

<table>
<tr><td>Jan</td><td>Feb</td><td>Mar</td><td>Apr</td><td>May</td><td>Jun</td><td>Jul</td><td>Aug</td><td>Sep</td><td>Oct</td><td>Nov</td><td>Dec</td></tr>
<tr><td colspan="3">Application Season</td><td colspan="8">Workshop Season</td><td rowspan="2">Summer Break</td></tr>
<tr><td colspan="3">Draw: April 1st</td><td colspan="8">Weekly meetings on Saturdays</td></tr>
</table>

2.2 Data

In this paper, I use a combination of administrative records, public available data on schools, and an endline survey.

Administrative Data. We use administrative data that contains student grades for multiple subjects from 2014 to 2020. Given that it was impossible to implement a name-to-name match with the official records (in compliance with the Peruvian Law that protects personal information), I use the following procedure to match our data. On one side, I got access to all student information including sex, birthday, and school identifiers. On the other side, from the program records, I also have the same variables. Given that the likelihood of matching gender and birthday in the same school is quite small, I assume that most matches are correct. Nevertheless, I only considered the cases where the probability of matching is 100 percent. With this conservative threshold, I was able to match approximately 70 percent of our sample. The administrative data contained enrollment status as well as final grades from most subjects. To be consistent across schools, I focus on the main subjects such as Spanish, Social Sciences, Religion, Arts, Mathematics, and Science. The Peruvian grading system goes from 0 to 20, and above 11 is considered a passing grade.

Endline Survey. I implemented an endline survey during the summer of 2020. I collected an online survey targeting all families who applied to the program in 2016. It is worth highlighting that the survey was conducted during the COVID-19 outbreak while Peru was in a very strict lockdown. Lima was particularly vulnerable because of its high density and a weak health system that was not ready to face the pandemic. The implementation of the survey was therefore challenging and I acknowledge that the results should be taken carefully given this unique situation. I designed a friendly online questionnaire for both girls and their parents. Families were able to fill out the survey from a desktop, laptop, smartphone, or tablet. They were reached via e-mail, text messages, and phone calls and our response rate was 20 percent approximately, a regular rate when it comes to online surveys. I collected information on 350 families during a month. The survey included two sections: one exclusively for parents and another one exclusively for applicants.[4] The survey sample was randomly selected from the program's administrative records, a pool of approximately 2000 applicants from 2016 to 2019. Families from older cohorts were harder to reach: phones and emails

[4] I am not able to verify if parents did not intervene when girls responded to their part of the survey. However, I constantly reminded them that they should leave girls to answer on their own. Results from the survey show that they indeed answered differently.

were not updated and some refused to fill the survey. Sample selection was stratified by application year and Appendix Table A.5 shows the balance between treatment and control in this sample. [5]

Our online questionnaire included the following sections:

1. *Perceptions on ability, grades, and effort.* I asked parents and applicants to rate their natural ability and academic achievement (grades) in four different subjects on a scale from 0 to 100. These subjects were Social Sciences, Spanish, Mathematics and Science.

2. *Time use.* I asked applicants to think about a typical day weekday and distribute 24 hours on different activities such as sleeping, school, eating, etc.

3. *Perceptions about college and future:* I asked applicants how happy they would be if they choose to follow a certain major (on a scale from 0-100). Majors included Journalism, Education, Law, Business, Sociology, Arts, Architecture, Mathematics, Medicine and Engineering.

4. *Aspirations and social norms:* I asked applicants to rate (on a scale from 0-100) some aspirations questions such as the major they would like to follow in college, the probability of finishing school with a high GPA, whether they talk to their parents about their educational future and which occupation they will like to have when they work. I also recreated a couple of social norms vignettes telling the hypothetical case of two women. One question asked if they agreed with following family advice (instead of their call) and the second one asked if they agreed with choosing a lady-like profession and marriage over professional success.

5. *Mental Health and Behavioral Outcomes:* I implemented a simple grit and mental health questionnaire. The grit questionnaire measures if girls can maintain focus, interest, and perseverance in obtaining their long-term goals while the mental health questionnaire measures signs of depression. I complement these questionnaires with a couple of questions to parents regarding how girls were dealing with the pandemic crisis.

[5] Table Appendix A.5 also shows the balance controlled by the year of application.

For simplicity, I constructed indexes for most of these outcomes. All our them are standardized for easier interpretation. Finally, I complement our analysis using publicly available data on school characteristics. I was able to match all applicants to their schools in the year they applied to the program I use this data for the randomization validation and selection analysis.

2.3 Selection

In this subsection, I examine the selection for applying to the program. First, I take a look at the observable characteristics of the pool of applicants from 2016 to 2019. The average applicant age is 9.4 for a range of 8-11 years old. Approximately, 90% of them reside in the wider Lima metropolitan area.[6]. They mostly come from the north and east side of Lima, as seen in Appendix Figure A.9. These districts are the most populated areas and have a high concentration of low and middle class families. Around 20% of the applicants come from public schools and 9% come from only-girls schools. The low amount of applicants coming from public schools might suggest a bias towards richer households except the Peruvian context is quite different. Approximately 50% of students are enrolled in private schools in Lima, and there is a lot of variation across prices. As in many developing countries, rates of privatization are even larger in urban areas. In Lima, the share of private schools increased from 23% in 2000 to 51% in 2017 (Allende, 2020).

Next, I use publicly available data from all primary schools in Lima and compare them with the schools where these applicants come from. I use information about the schools' neighborhoods and I measure the percentage of poverty surrounding the school and private school fees as a proxy for families' socioeconomic background. As seen in Appendix Figure A.8, there are no systematical differences between applicants' schools and the universe of schools in Lima. This suggests that in terms of socioeconomic background, our sample is representative on average of the population, with most applicants coming from middle-class neighborhoods, as suggested before, or more explicitly, most applicants' schools are in

[6]Some parents outside the city applied hoping to get the scholarship and planned to move or commute to the city during the weekends.

neighborhoods with a low share of poor households. What is more, applicants coming from private schools pay approximately 100 USD per month in school fees, which is equivalent to one-third of Peru's minimum monthly wage at that time.

I also compare the test scores results using the ECE (*Examen Censal de Estudiantes*), a national standardized test for second graders, as seen in Figure 1. Most applicants come from schools with significantly higher performance on these tests, both in Mathematics and Spanish. This suggests that applicants do come from higher-quality schools, implying that parents invest more in their education, regardless of their economic status. Overall, I see that there is selection on who applies to the program with a clear profile: applicants come from families that are from the middle class and attend schools with better education performance than their eligible peers. In addition, during the 2019 application period, we collected a baseline survey and confirmed that most of the applicants' parents have a college degree or more.[7]

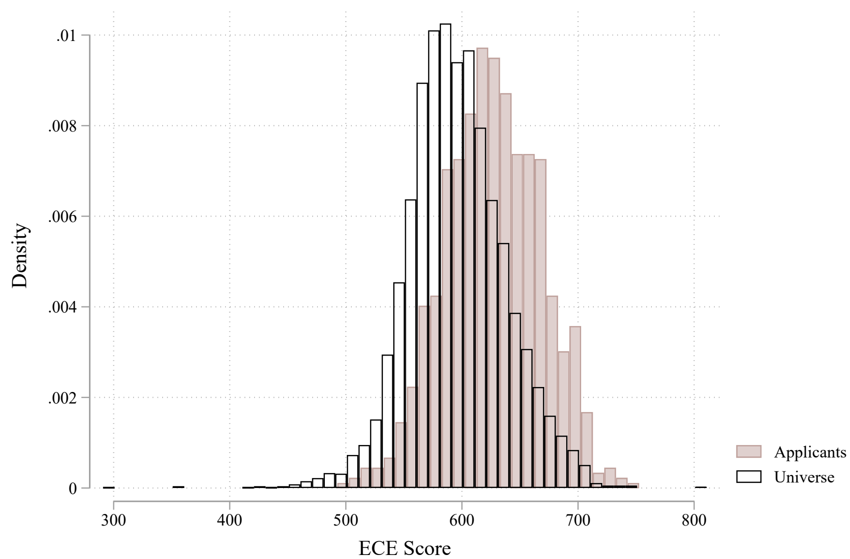
To understand some of our academic achievement results, it is important to see the relative ranking of the applicant compared with their school cohort. Using the matched data with the administrative records (as detailed in the previous section), I compare the pool of applicants with students in the same class the year before applying to the program. In Figure 2 I see that applicants to the program have higher grades in STEM courses. The distribution of applicants seems to be bi-modal showing a small group of applicants having significantly higher grades. The results are quite similar when looking at non-STEM courses as seen in Appendix Figure A.10. I pay a special look at this since it shows how much room these girls have to improve their grades. Given that these girls were already good students, they had limited space to improve their academic performance. [8]

[7]More details about the baseline are described in Appendix A.1

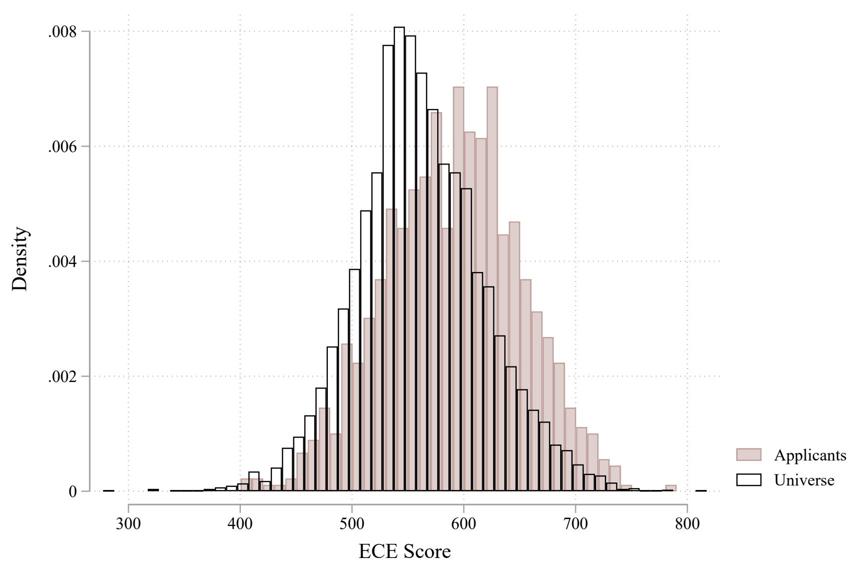
[8]For all of these grade distributions, I have that their Kolmogorov-Smirnov test rejects the equality of the two distributions with $pvalue \leq 0.001$.

Figure 1: School Average on National Standardized Test for 2nd graders

(a) Spanish



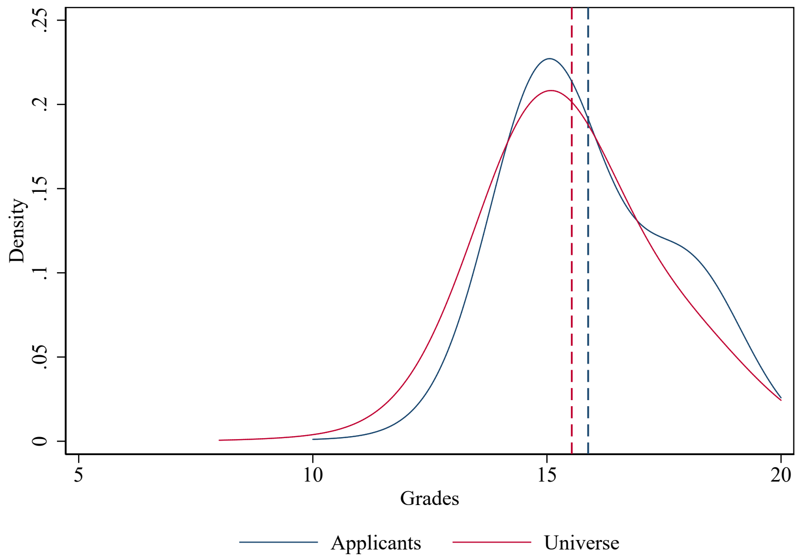
(b) Mathematics



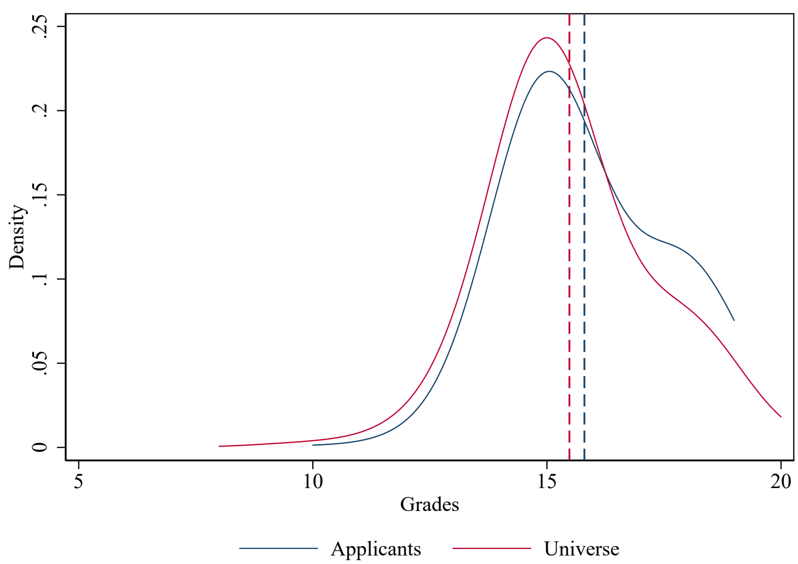
Source: 2014 ECE average school test scores. The universe represents all schools in the metropolitan area of Lima.

Figure 2: Selection in School STEM Subjects Performance

(a) Science



(b) Mathematics



Source: I use grades measured at the end of the academic year prior to the participation year. The universe represents all applicants' peers in the same school and classroom.

3 Empirical Strategy

Every year, the program randomly selects new scholars in a lottery transmitted live on social media. After being selected, parents have a few days to accept or not the scholarship. They have to sign a commitment to bring their kid every Saturday for the weekly workshops. After 3 unjustified missed classes, they are removed from the program. There were only a few anecdotal cases where people rejected the scholarship and were automatically replaced with another randomly selected applicant before the program started. To validate the empirical strategy, I analyzed the randomization implemented by the NGO. I find that conditional on applying, there are no systematic differences between applicants who obtained the scholarship ("treated" students) and those who did not ("control" students) as seen in Table 2. What is more, by looking at the geographical distribution of applicants in Appendix Figure A.11 (Address GPS location at application time), treated applicants' households are similarly dispersed around the city to their control counterparts.

Table 2: Balance between control and treatment groups

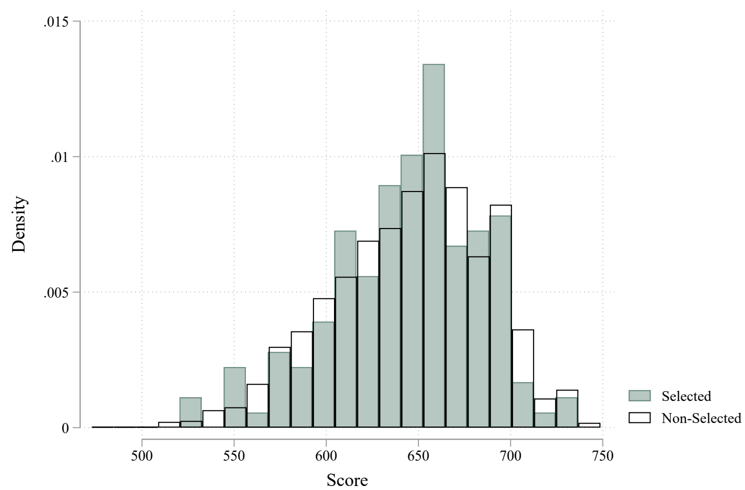
	(1)	(2)	(3)
Variable	Control	Treatment	Difference
Age	9.465	9.456	-0.009
	(1.253)	(1.036)	(0.101)
Grade	4.486	4.419	-0.067
	(1.325)	(1.200)	(0.108)
Public	0.194	0.196	0.002
	(0.396)	(0.398)	(0.033)
Poverty Around School	0.159	0.146	-0.013
	(0.132)	(0.116)	(0.011)
School Fee (PEN)	677.7	721.0	43.32
	(543.7)	(583.9)	(50.21)
ECE Spanish	645.9	643.9	-1.958
	(42.70)	(41.10)	(3.601)
ECE Math	611.6	608.4	-3.262
	(58.49)	(56.05)	(4.931)
No. Teachers	22.23	24.17	1.94
	(15.39)	(17.77)	(1.276)
No. Students	412.6	428.5	15.96
	(295.71)	(339.74)	(24.51)
Observations	2,529	160	2,691

Note: Poverty around school is measured by the fraction of poor households within 3 km. School fees is for those attending private schools only. All school variables reported are from the 2014 School Census.

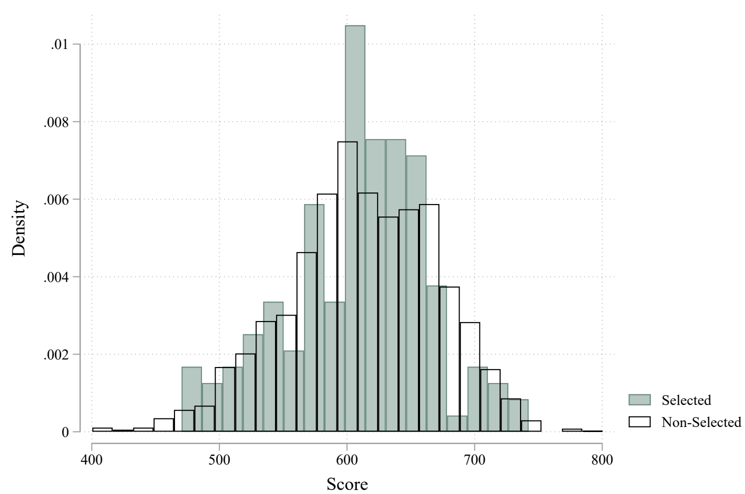
Using the same measures as in section 2.3 (the percentage of extremely poor households around schools), treated and control students are similar in terms of their neighborhood poverty distribution as well as private school fees, as seen in Figure A.12. Additionally, I do not find significant differences in the quality of schools attended by treated and control students as measured by average test scores for 2nd graders, both for mathematics and Spanish, as seen in Figure 3. All this evidence suggests that conditional on applying to the program, applicants are selected randomly.

Figure 3: School Average on National Standardized Test for 2nd graders

(a) Spanish



(b) Mathematics



Source: 2014 ECE average school test scores.

The empirical design relies on the lottery implemented by the NGO each year. As shown before, conditional on applying, the selection of participants is random. Therefore, this paper leverages on the random assignment to the scholarship and I use two main sources of data detailed in the data section to study the outcomes of interest. I pool the data from every cohort since 2016 to 2019 and estimate the following equation:

$$y_{it} = \beta_{it} + \delta X_{it} + \phi_t + \epsilon_{it} \tag{1}$$

where y_{it} are the outcomes for each applicant i of a lottery cohort t, and t corresponds to years 2016 to 2019. $Treat_{it}$ is an indicator variable that equals to one if an applicant is randomly selected to participate in the program and β captures the treatment effect of the program. X_{it} are control variables (for precision) that include parent's education level (or school fee as a socioeconomic proxy when using the administrative data) and applicants' age, and ϕ_t are year fixed effects. Additionally, for the academic achievement results, I control for baseline grades. These estimations have robust standard errors using the Huber-White/sandwich estimator.

4 Results

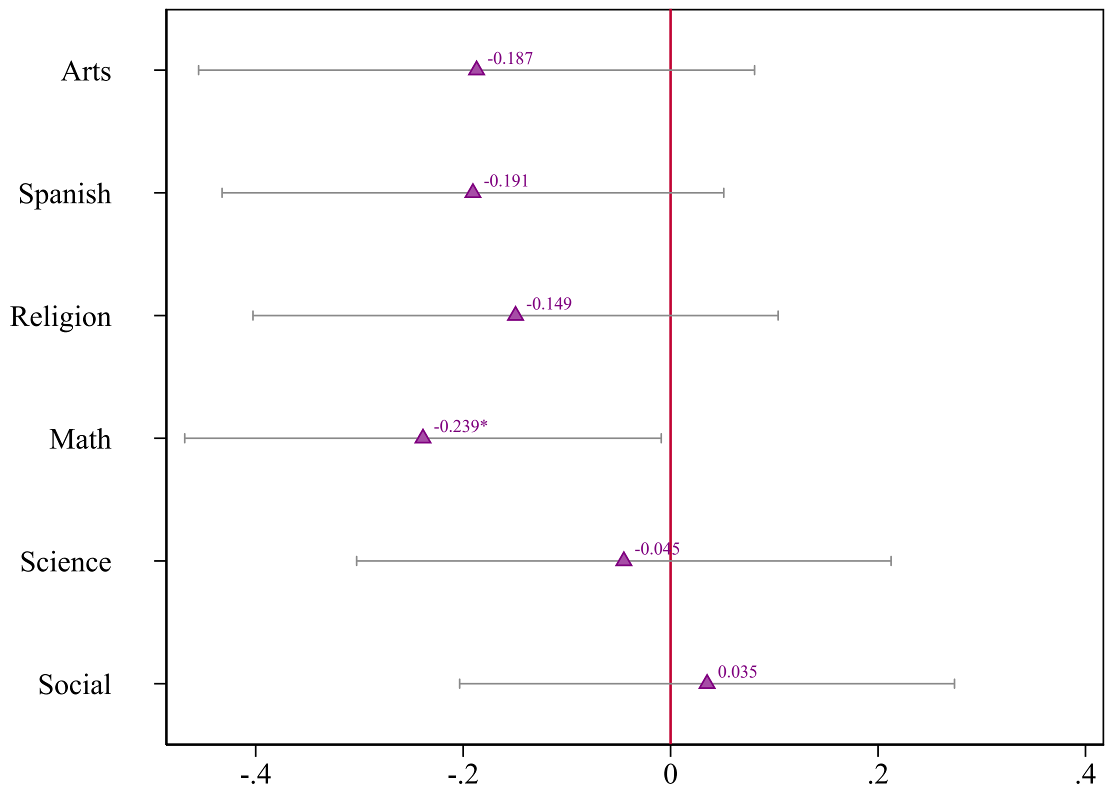
4.1 Academic Achievement

First, I study the effects of the program on academic performance. I use administrative data from the Ministry of Education collected from every school at the end of the academic year. I estimate Equation (1) including a baseline control of academic performance the year before the intervention started. As seen on 4, I find slightly negative effects on Mathematics, Religion, Spanish and Arts and nearly zero effects on Science and Social Sciences. Figure 5 shows the same results for the year after they are still looking non-significant. If anything, there is some evidence of a "catching-up" effect that continues after 2 years of participating in the program as seen in Figure 6. [9]. Overall, given the lack of significance in the estimates,

[9] Tables for each one of the figures are including on Appendix Tables A.6 ,A.7, A.8

the results suggest that the program had no impact on academic performance in the short-term. I highlight that the program itself was not remedial and in fact, it had no relationship with the school curricula. Thus the only possible channel where the program could have any effects on grades was through behavioral changes, which I consider as the first stage and it discussed these results in the next section.

Figure 4: Effects on academic performance during the treatment year



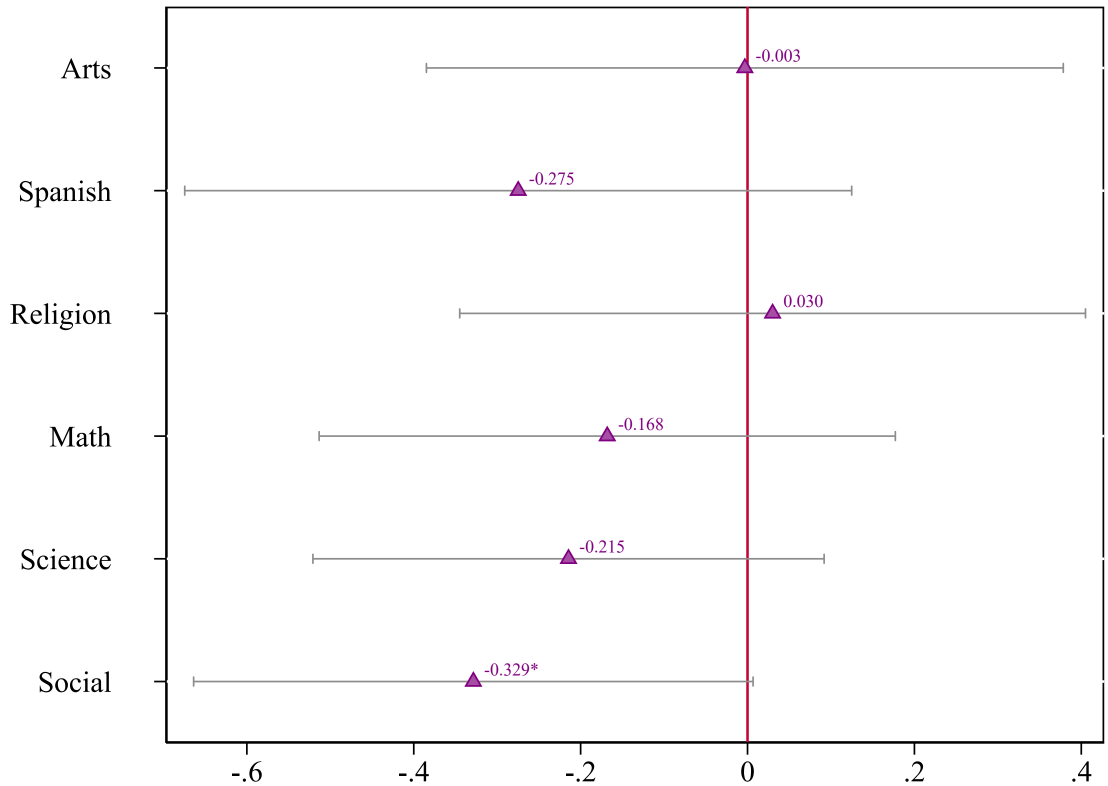
Source: administrative school records. I use grades reported at the end of the academic year. Notes: Treatment effect coefficient are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, school fees as a proxy for SES, and baseline grades. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

4.2 Behavioral Responses

This section presents the results using the endline survey described in Section 2.3.

Perceptions on Ability and Grades: I asked parents and applicants to rate their

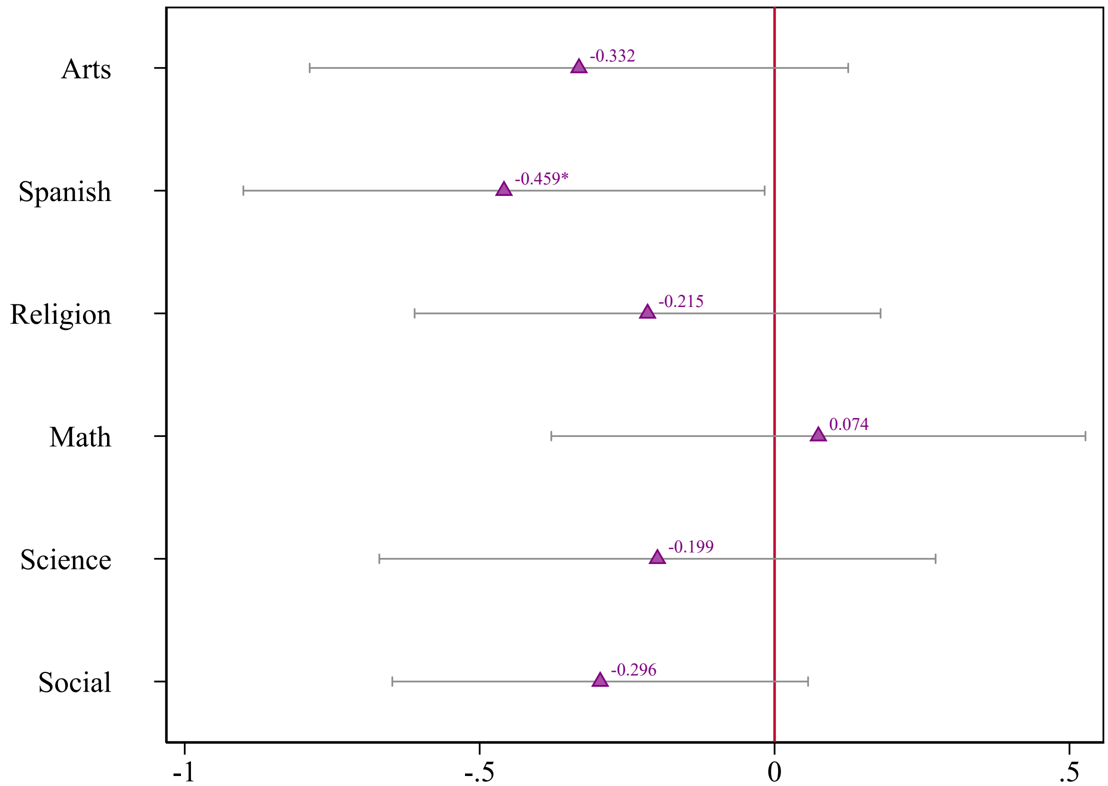
Figure 5: Effects on academic performance after 1 year



Source: administrative school records. I use grades reported at the end of the academic year. Notes: Treatment effect coefficients are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, school fees as a proxy for SES, and baseline grades. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

ability and achievement (grades) in different subjects on a scale from 0 to 100. I find no significant effects on applicants' perceptions of ability, both on STEM and non-STEM subjects as seen in Figure 7 (a). However, when looking at perceptions of grades, I see that girls in the treatment group are more confident about their grades in all subjects except Spanish. In particular, I find significant effects: a 5 percent increase relative to the control group, on Science as seen in Figure 7 (b). I do not find significant effects on parents' perceptions nor in differences between parents and applicants' responses. Overall, I see evidence suggesting that girls who participate in the program tend to be more confident about their grades relative to their parents' perceptions as seen in Appendix Figure A.16. I also asked both parents and

Figure 6: Effects con academic performance after 2 years



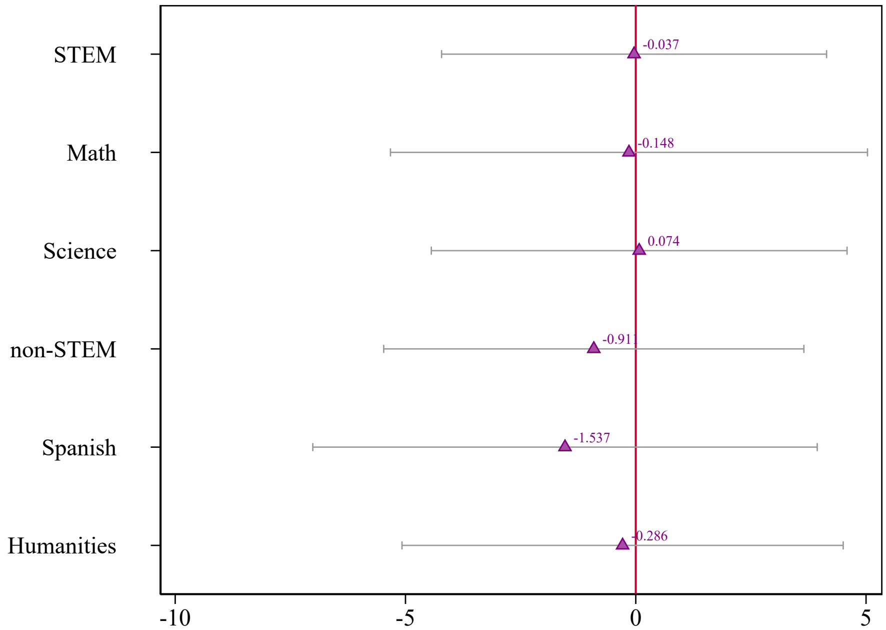
Source: administrative school records. I use grades reported at the end of the academic year. Notes: Treatment effect coefficients are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, school fees as a proxy for SES, and baseline grades. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

applicants about their effort at school and I find no significant effects albeit negative as seen on Table 3.

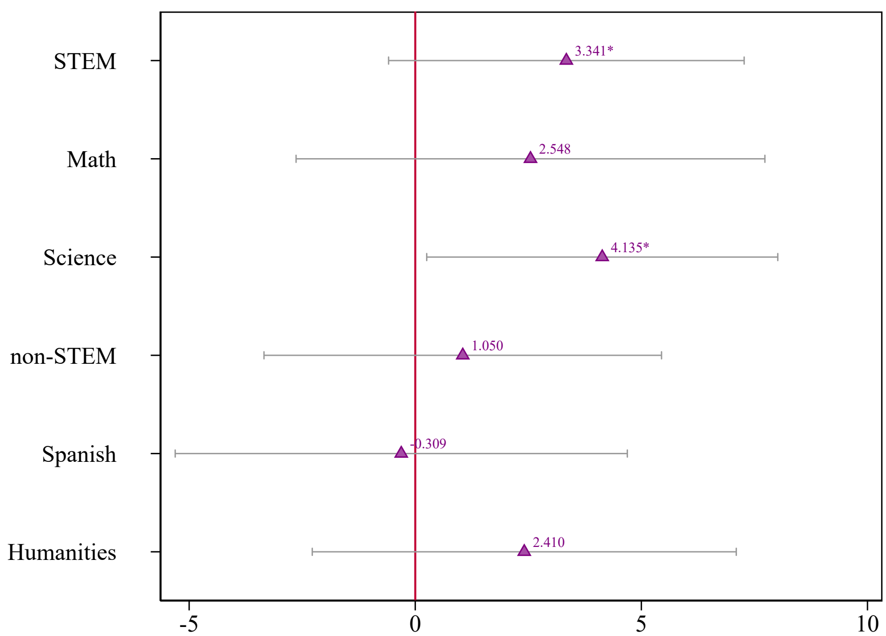
Overconfidence in Academic Performance: As described before, the pool of applicants had higher grades in school compared to their classroom peers. Using a gaussian density to plot the difference between real and perceptions of academic performance calculated using deciles of the school-cohort distribution, I find no statistical difference between the treatment and control groups. Figure 8 shows the distribution of both treatment and control in STEM subjects (Mathematics and Science). Notably, I see that the program had higher effects on those students with the highest level of beliefs distortion and vice versa. This is not the case

Figure 7: Effects on perceptions of ability and grades by subject

(a) Ability



(b) Academic Performance (grades)



Note: Data sources are endline data. Treatment effect coefficients are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, parents' education as a proxy for SES, a dummy that indicates if the girl applied more than once to the program. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

Table 3: Effects on effort at school

	Parent		Applicant		Difference	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-1.716	-1.875	-0.525	-0.866	-1.531	-1.311
	(1.833)	(1.908)	(2.114)	(2.191)	(1.439)	(1.480)
Controls	No	Yes	No	Yes	No	Yes
Mean	88.63	88.63	85.90	85.90	2.878	2.878
Obs.	350	349	338	337	338	337

All models have year FE. Controls: parents; education , applicant's age, and a dummy that indicates whether the girl applied more than once. Significance: *** p<0.01 ** p<0.05 * p<0.1.

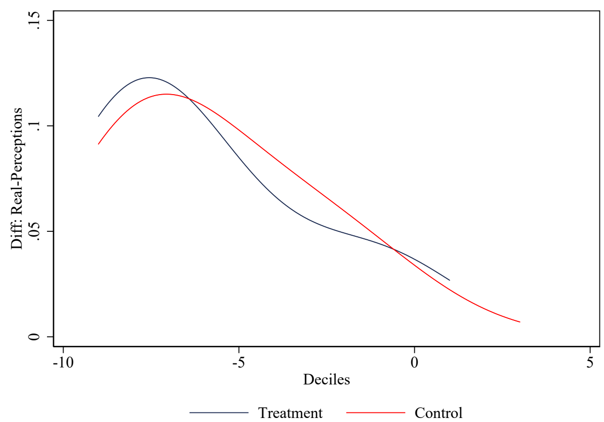
when I look at non-STEM subjects as in Appendix Figure A.18.

Time Use Effects: Figure 9 shows the effects on time use reported by applicants when I asked about how a typical day (24 hours) looks like for them. I find significant positive effects on time spent on Personal Projects, which is consistent with some qualitative interviews I had with the NGO staff suggesting that participants reported starting their own science projects after concluding the program. The effect is 30 a percent increase that is equivalent to 10 more minutes per day on average. Also, I see significant negative effects on time spent on homework. Girls spend on average 9 percent less time relative to the control group on homework and schooling, which is approximately 25 minutes per day. Additionally, I find negative effects on sleeping but these results are not significant. The findings suggest that girls in the treatment group might be more productive (as measured by their perceptions of their grades), use less time in school and trade it off with their own personal projects, which is consistent with the negative effects I see regarding their effort at school. Table A.15 in Appendix shows the results with no controls, and the results are consistent as well.

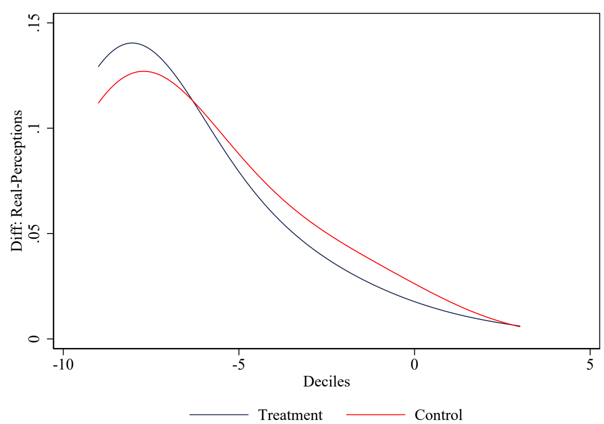
Perceptions about Majors in College: I ask how happy the applicant can be if they choose to follow a certain major (on a scale from 0-100). Figure 10 (b) shows a very clear path: there are negative significant effects on non-STEM careers (14 percent decrease relative to control group) mainly driven by majors like Law (24 percent decrease), Journalism (20 percent decrease) and Education (12 percent decrease) for treated girls. Surprisingly, in Figure 10 (b) I see small negative significant effects on Math and Architecture majors and

Figure 8: Effects on (over) confidence in STEM grades

(a) Math

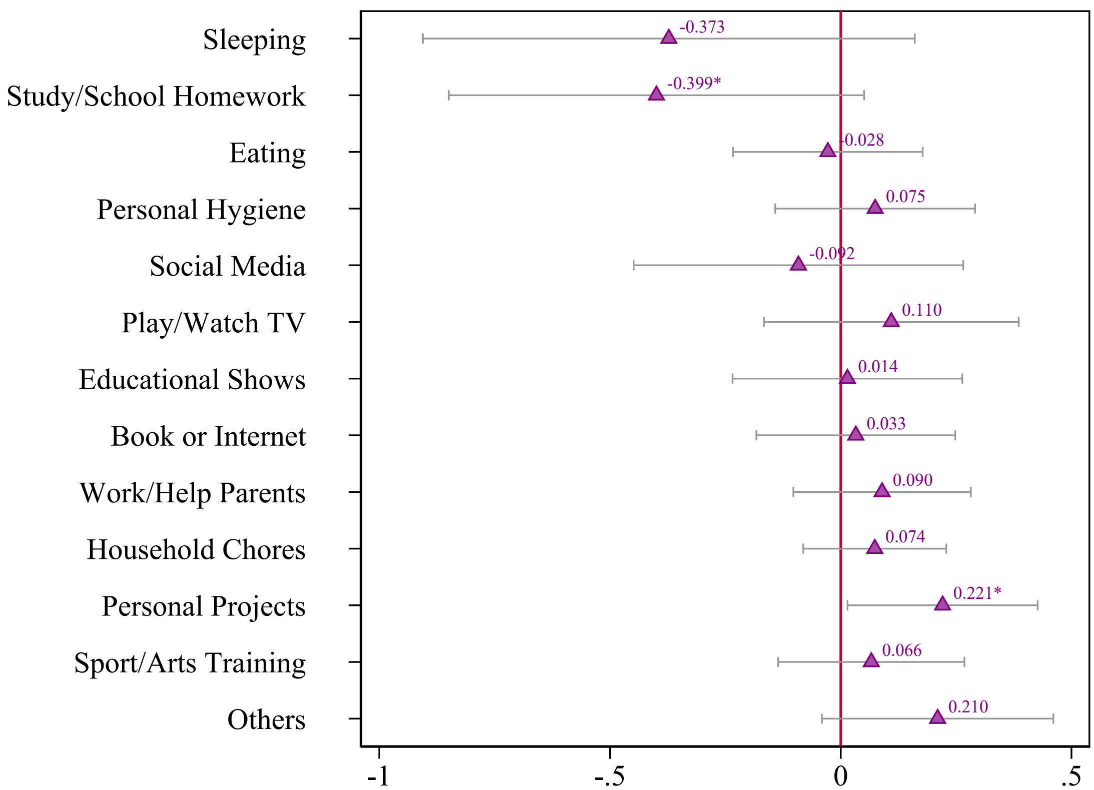


(b) Science



Note: Data sources are endline and administrative data. Deciles are calculated at the school-cohort level. Densities are estimated using a Gaussian function.

Figure 9: Effects on time use (hours)



Note: Data sources are endline data. Treatment effect coefficients are reported for each activity. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, parents' education as a proxy for SES, a dummy that indicates if the girl applied more than once to the program. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

no significant negative effects on Medicine and Engineering. It is important to highlight that the program workshops were focused on topics related to biology and physics, and not mathematics. Given the selection into applying to the program that I documented before and that applicants have a high interest in science to begin with, we should take these results carefully.[10] On average, girls who applied to the program were already very enthusiastic about STEM careers. However, being exposed to this program and top scientists might have also made these girls more realistic about their chances of succeed at any career and decrease their overall confidence. In this sense, I find that treated girls seem to be less pessimistic about Engineering and Medicine majors in comparison with other fields.[11]

The previous evidence is consistent with another measure: I asked applicants about their perceptions related to happiness for different major choices but for typical girls (same age and neighborhood). As seen in Figure 11 (b), there is clear evidence of negative effects on STEM majors, following the same pattern as their own perceptions. Nevertheless, when looking at Non-STEM majors, treated girls are more pessimistic about their peers' likelihood to be happy on Education, Sociology or Arts. These graphs confirm the idea that overall, treated girls seems to be more pessimistic about the likelihood of being happy and succeeding in any career, not only for themselves but also for girls like them.

Aspirations, Social Norms, Mental Health and Behavioral Outcomes: Inspired by Jayachandran (2015), I created an aspiration index, including 4 measures: probability of finishing high school with an outstanding grade, whether they talk with their parents about their future education, progressiveness of the occupation they will like to work on, and whether their major preference is STEM. I do not find significant effects on this index as seen on Table 4.

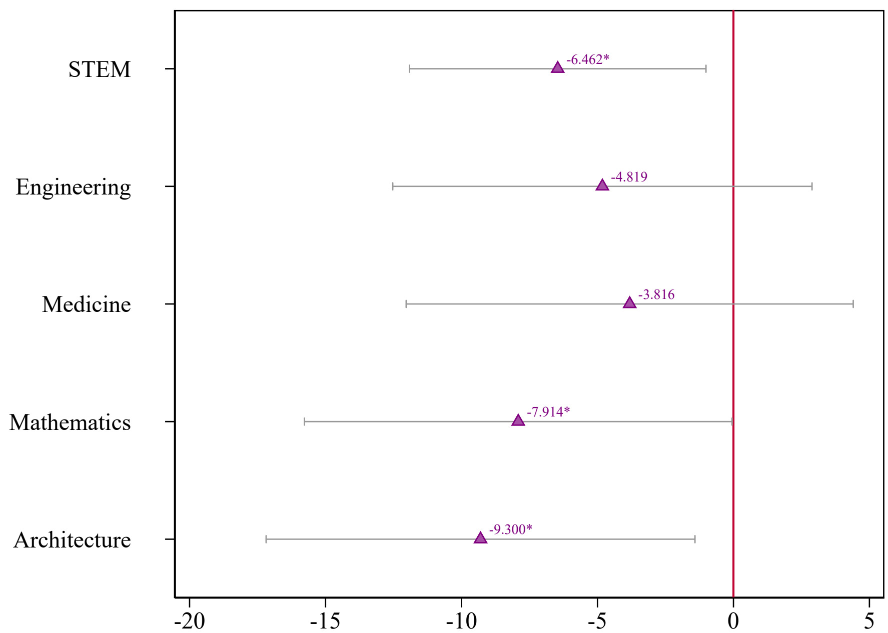
It is worth highlighting that there are significant negative effects on the likelihood of talking to their parents about their future. According to NGO staff, most girls developed strength and independence as they were constantly challenged with weekly homework that

[10]For example, I do not find significant effects on their preferred major. Only a small positive but not significant effect on health-related majors. See Table B.9 and B.10.

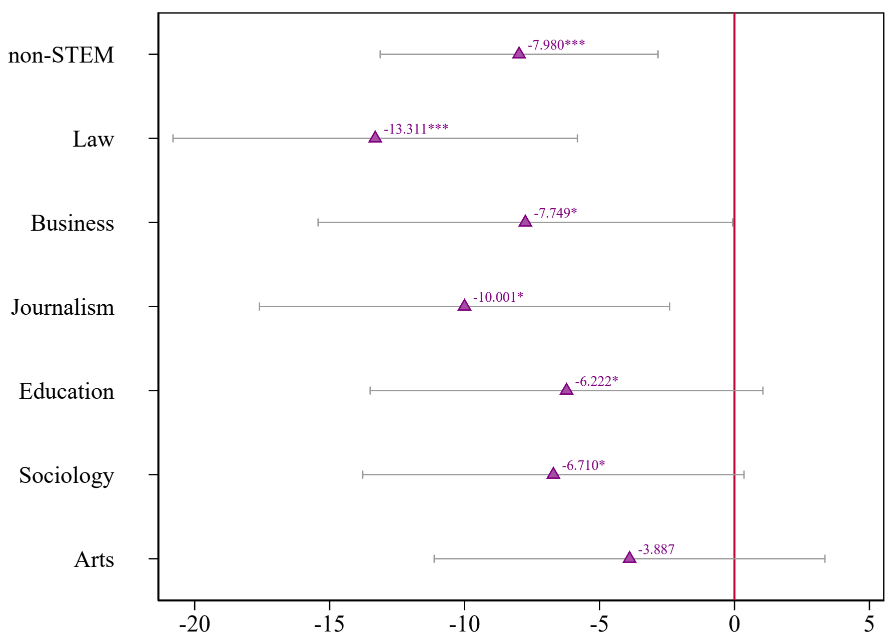
[11]I also do not find an effect when I asked the same question to parents, as seen on Appendix Tables B.2 and Table B.6.

Figure 10: Effects on perceptions about majors in college

(a) STEM



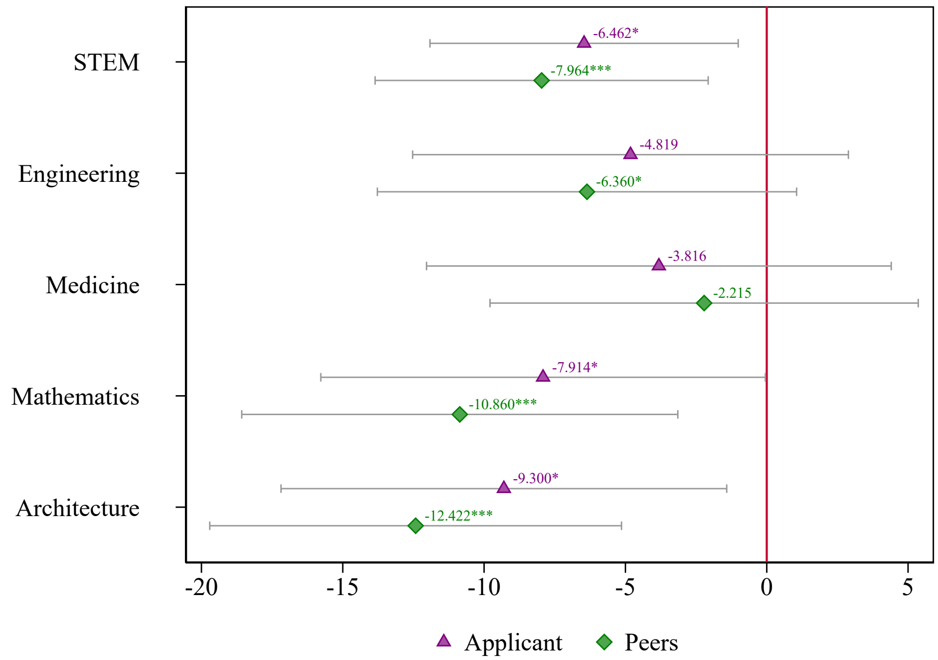
(b) Non-STEM



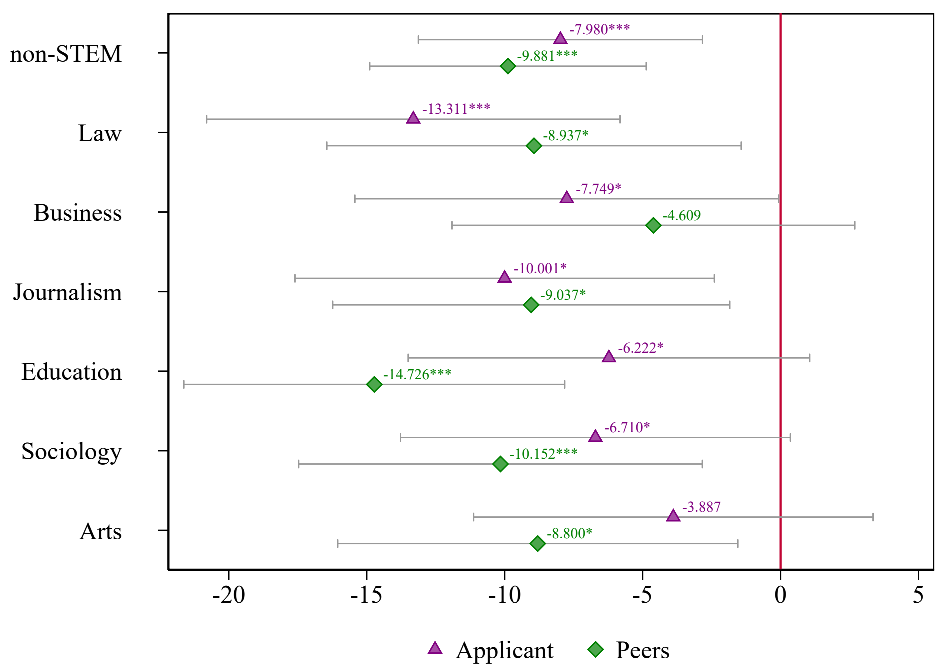
Note: Data sources are endline data. Treatment effect coefficients are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, parents' education as a proxy for SES, a dummy that indicates if the girl applied more than once to the program. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

Figure 11: Effects on perceptions about majors in college for applicants and their peers

(a) STEM



(b) Non-STEM



Note: Data sources are endline data. Treatment effect coefficients are reported for each subject. Linear regressions are estimated as detailed in Section 3.2. Controls include year fixed effects, age, parents' education as a proxy for SES, a dummy that indicates if the girl applied more than once to the program. Standard errors are robust. The gray lines are 95 percent confidence intervals. Significance: $***p < 00.01$, $**p < 0.05$, $*p < 0.01$

Table 4: Effects on aspirations (index)

	Finishing HS with 15+ GPA		Talked to parents about future		Progressiveness of occupation		Major preference is STEM		Aspiration Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	0.006	0.018	-0.352**	-0.334*	0.100	0.087	-0.005	-0.007	-0.033	-0.027
	(0.122)	(0.122)	(0.170)	(0.173)	(0.120)	(0.122)	(0.127)	(0.129)	(0.083)	(0.084)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Obs.	338	337	330	329	309	308	315	314	298	297

Progressiveness of occupation is the average of Male-dominated status that preference has, which is, in turn, measured as the difference between the percentage of males and females in an occupation.
All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table 5: Effects on social norms (vignettes)

	Follows family advice		Marriage and lady-like career	
	(1)	(2)	(3)	(4)
Treated	-5.831**	-6.068**	0.087	-0.385
	(2.425)	(2.439)	(2.398)	(2.474)
Controls	No	Yes	No	Yes
Mean	19.46	19.46	11.54	11.54
Obs.	330	329	327	326

All models have year FE. Controls include parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1.

demanded public presentations and intense debates with lecturers. In some anecdotal conversations, some staff members mentioned that girls were also more intense in their manners and outspoken compared to girls of the same age. Not surprisingly, when looking at the results on social norms, I also see that they are less likely to take and follow family advice regarding their future as seen in Table 5.

I also do not find significant effects on social norms regarding marriage and choosing a lady-like career, which might be consistent with the fact that most girls who applied to the program were already into STEM in the first place. I also see that girls in the treatment group show less grittiness and this is consistent across all components of this index, as seen in Table 8. Notably, I believe that the pandemic could have exacerbated some of these traits and girls might feel worse off given their high expectations and the limitations imposed by the strict lockdown.

Because this survey was implemented during the COVID-19 pandemic, I also asked some questions about their emotional state and how they were facing those stressful times. Table 6 shows a mental health index, that sheds a light on how girls see themselves. Girls in the treated group reported feeling like they do not have many qualities, feeling less calm, feeling less likely to do things well and more stressed, but they do feel happy with themselves and full of energy. Their parents also reported that during the pandemic (and remote schooling) they were able to keep up with the school homework, but social distance was affecting them. Even when these results are not significant, I take them as a sign that girls in the treatment group were negatively affected by the pandemic in terms of their mental health and emotions.

Table 6: Effects on mental health (index)

	Happy with myself (1)	I have many qualitites (2)	I do things well (3)	I feel calm (4)	Not stressed (5)	I feel full of energy (6)	Mental Index (7)
Treated	0.119	-0.087	-0.140	-0.017	-0.011	0.021	-0.019
	(0.124)	(0.142)	(0.142)	(0.133)	(0.122)	(0.123)	(0.095)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	326	326	326	326	326	326	326

A higher score in mental index's individual components means better mental health.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
All models have year FE.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table 7: Effects on COVID-19 emotional state

	Discipline in school		Social distancing		Covid Coping Index	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.039	0.043	-0.052	-0.053	-0.007	-0.005
	(0.121)	(0.121)	(0.119)	(0.119)	(0.111)	(0.111)
Controls	No	Yes	No	Yes	No	Yes
Obs.	346	345	346	345	346	345

All models have year FE. Controls include parents' education , applicant's age, and a dummy that indicates whether the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1.

Table 8: Effects on grittiness (index)

	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7	Component 8	Grit Index
Treated	-0.204	-0.323**	-0.204	-0.148	-0.157	-0.059	-0.314**	-0.350**	-0.220**
	(0.185)	(0.157)	(0.183)	(0.159)	(0.184)	(0.180)	(0.150)	(0.147)	(0.104)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	3.482	4.018	3.928	4.189	3.761	3.486	4.288	4.369	3.940
Obs.	326	326	326	326	326	326	326	326	326

Component 1: New ideas and projects do not distract me from previous ones. Component 2: Setbacks don't discourage me. Component 3: I have been obsessed with a certain idea or project without losing interest afterwards. Component 4: I am a hard worker. Component 5: I often set a goal and I do not stop until I complete it. Component 6: I do not have difficulty maintaining my focus on projects that take more than a few months to complete. Component 7: I finish whatever I begin. Component 8: I am diligent and I have initiative. A higher score in grit index's individual components means having more grit. All models have year FE. Significance: *** $p<0.01$ ** $p<0.05$ * $p<0.1$.

5 Conclusions

In this paper, I study a novel private initiative that highlights a crucial issue related to gender inequality: the low participation of females in STEM fields in a country grappling with severe gender issues. According to Plan International (2018), Lima (Peru) is one of the worst cities globally for women and girls. Many of the gender issues in Peru stem from entrenched social norms and stereotypes that impact individuals from an early age. This paper examines a program designed to empower young girls, boosting their self-esteem and confidence in STEM—a field heavily dominated by men—in a context where gender violence is prevalent. Importantly, the program operates within a girls-only safe environment, encouraging them to develop skills like public speaking, debating, and questioning without the pressure of interacting with boys.

This study contributes new and valuable insights to the existing literature. To my knowledge, few studies focus on STEM aspirations at such a young age, with most interventions targeting adolescents in secondary school or recent high school graduates. By analyzing the impacts and limitations of this program, I aim to inform future scale-ups and adaptations in similar settings. Notably, the results suggest that the girls who applied were already high achievers, which leaves limited room for academic improvement.

The short-term results presented in this paper offer a clearer understanding of the program's mechanisms and its limitations. By utilizing the lottery system used for participant selection, I provide evidence on both subjective (aspirations and beliefs) and objective (grades) outcomes to assess the program's effects. I find no evidence of improvements in academic performance, as measured by grades, up to two years after participating in the program. However, there were indications of increased confidence in science performance (objective) but no changes in effort or perceptions of ability (subjective). What is more, girls who participated in the program seem to be more productive as those in the treatment group, who were already top performers at school, seem to have shifted focus to other non-school activities without impacting their grades. This suggests they are managing their time effectively. Additionally, the increased independence reported by participants—such as talking less to their parents or being less likely to follow family advice—aligns with these findings. On the other hand, being more productive or having more to do might also be affecting their grittiness and mental health, specially during stressful periods such a the COVID-19 pandemic. Further research into longer-term outcomes would be valuable to fully understand these dynamics.

References

Martha J Bailey and Susan M Dynarski. Gains and gaps: Changing inequality in u.s. college entry and completion. Working Paper 17633, National Bureau of Economic Research, December 2011.

Mark Rosenzweig and Junsen Zhang. Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of chinese twins. *Journal of Development Economics*, 104(C):245–260, 2013.

Glenn Ellison and Ashley Swanson. The gender gap in secondary school mathematics at high achievement levels: Evidence from the american mathematics competitions. *Journal of Economic Perspectives*, 24(2):109–28, June 2010. doi: 10.1257/jep.24.2.109.

S. Cheryan, A. Master, and A. N. Meltzoff. Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology*, 6(49), February 2015.

Ainara Gonzalez De San Roman and Sara De La Rica. Gender gaps in pisa test scores: The impact of social norms and the mother's transmission of role attitudes. *Estudios de Economia Aplicada*, 34:79–108, January 2016.

Alex Eble and Feng Hu. Gendered beliefs about mathematics ability transmit across generations through children's peers. *Nature Human Behaviour*, 2022. doi: https://doi.org/10.1038/s41562-022-01331-9.

Seema Jayachandran. The roots of gender inequality in developing countries. *Annual Review of Economics*, 7(1):63–88, 2015.

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

Sarah R Cohodes, Helen Ho, and Silvia C Robles. Stem summer programs for underrepresented youth increase stem degrees. Working Paper 30227, National Bureau of Economic Research, July 2022. URL http://www.nber.org/papers/w30227.

Lucia Del Carpio and Maria Guadalupe. More women in tech? evidence from a field experiment addressing social identity. *Management Science*, 68(5):3196–3218, October 2021.

Joseph A. Kitchen, Gerhard Sonnert, and Philip M. Sadler. The impact of college- and university-run high school summer programs on students' end of high school stem career aspirations. *Science Education*, 102(3):529–547, 2018. doi: https://doi.org/10.1002/sce.21332. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.21332.

Corinne Moss-Racusin, Evava Pietri, Erin Hennes, John Dovidio, Victoria Brescoll, Gina Roussos, and Jo Handelsman. Reducing stem gender bias with vids (video interventions for diversity in stem). *Journal of Experimental Psychology. Applied*, 24, May 2018. doi: 10.1037/xap0000144.

Sule Alan and Seda Ertac. Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy*, 126(5):1865 – 1911, 2018.

Oriana Bandiera, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. Women's empowerment in action: Evidence from a randomized control trial in africa. *American Economic Journal: Applied Economics*, December (1):210–59, January 2020. doi: 10.1257/app.20170416.

Diva Dhar, Tarun Jain, and Seema Jayachandran. Intergenerational transmission of gender attitudes: Evidence from india. *The Journal of Development Studies*, 55(12):2572–2592, 2019. doi: 10.1080/00220388.2018.1520214.

Diva Dhar, Tarun Jain, and Seema Jayachandran. Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, 112 (3):899–927, March 2022. doi: 10.1257/aer.20201112. URL https://www.aeaweb.org/articles?id=10.1257/aer.20201112.

James J. Heckman and Yona Rubinstein. The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, 91(2):145–149, May 2001. doi: 10.1257/aer.91.2.145.

Lori Beaman, Esther Duflo, Rohini Pande, and Petia Topalova. Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science (New York, N.Y.)*, 335:582–6, February 2012. doi: 10.1126/science.1212382.

Bernard Tanguy, Stefan Dercon, Kate Orkin, and Alemayehu Seyoum Taffesse. The future in mind: Aspirations and forward-looking behaviour in rural ethiopia. Cepr discussion paper no. dp10224, October 2014.

Thomas Breda, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools. *The Economic Journal*, 133(653):1773–1809, 02 2023. ISSN 0013-0133. doi: 10.1093/ej/uead019. URL https://doi.org/10.1093/ej/uead019.

Tara Dennehy and Nilanjana Dasgupta. Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences of the United States of America*, 114, May 2017. doi: 10.1073/pnas.1613117114.

Melissa S Kearney and Phillip B Levine. Role models, mentors, and media influences. *The Future of Children*, 30(2020):83–106, 2020.

Catherine Porter and Danila Serra. Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12:226–254, July 2020. doi: 10.1257/app.20180426.

Alex Eble and Feng Hu. Child beliefs, societal beliefs, and teacher-student identity match. *Economics of Education Review*, 77:101994, 2020. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2020.101994.

Scott E. Carrell, Marianne E. Page, and James E. West. Sex and Science: How Professor Gender Perpetuates the Gender Gap*. *The Quarterly Journal of Economics*, 125(3):1101–1144, 08 2010.

Amanda L. Griffith. Faculty gender in the college classroom: Does it matter for achievement and major choice? *Southern Economic Journal*, 81(1):211–231, 2014. doi: https://doi.org/10.4284/0038-4038-2012.100. URL https://onlinelibrary.wiley.com/doi/abs/10.4284/0038-4038-2012.100.

Amanda L. Griffith and Joyce B. Main. First impressions in the classroom: How do class characteristics affect student grades and majors? *Economics of Education Review*, 69: 125–137, 2019. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2019.02.001. URL https://www.sciencedirect.com/science/article/pii/S0272775718304151.

Amanda L. Griffith and Joyce B. Main. The role of the teaching assistant: Female role models in the classroom. *Economics of Education Review*, 85:102179, 2021. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2021.102179. URL https://www.sciencedirect.com/science/article/pii/S0272775721000972.

Michela Carlana. Implicit Stereotypes: Evidence from Teachers' Gender Bias*. *The Quarterly Journal of Economics*, 134(3):1163–1224, 03 2019.

George A. Akerlof and Rachel E. Kranton. Economics and Identity. *The Quarterly Journal of Economics*, 115(3):715–753, 08 2000.

Roland Benabou and Jean Tirole. Identity, Morals, and Taboos: Beliefs as Assets. *The Quarterly Journal of Economics*, 126(2):805–855, 05 2011.

Travis J. Lybbert and Bruce Wydick. Hope as Aspirations, Agency, and Pathways Poverty Dynamics and Microfinance in Oaxaca, Mexico. In *The Economics of Poverty Traps*. University of Chicago Press, 12 2018. ISBN 9780226574301. doi: 10.7208/chicago/9780226574448.003.0006. URL https://doi.org/10.7208/chicago/9780226574448.003.0006.

Fernando Saltiel. Multi-Dimensional Skills and Gender Differences in STEM Majors. *The Economic Journal*, 133(651):1217–1247, 10 2022. ISSN 0013-0133. doi: 10.1093/ej/ueac079. URL https://doi.org/10.1093/ej/ueac079.

Maria Micaela Sviatschi and Iva Trako. Gender violence, enforcement, and human capital: Evidence from women's justice centers in peru. *Journal of Development Economics*, 168: 103262, 2024. ISSN 0304-3878. doi: https://doi.org/10.1016/j.jdeveco.2024.103262. URL https://www.sciencedirect.com/science/article/pii/S0304387824000117.

A Appendix

A.1 2019 Baseline Results

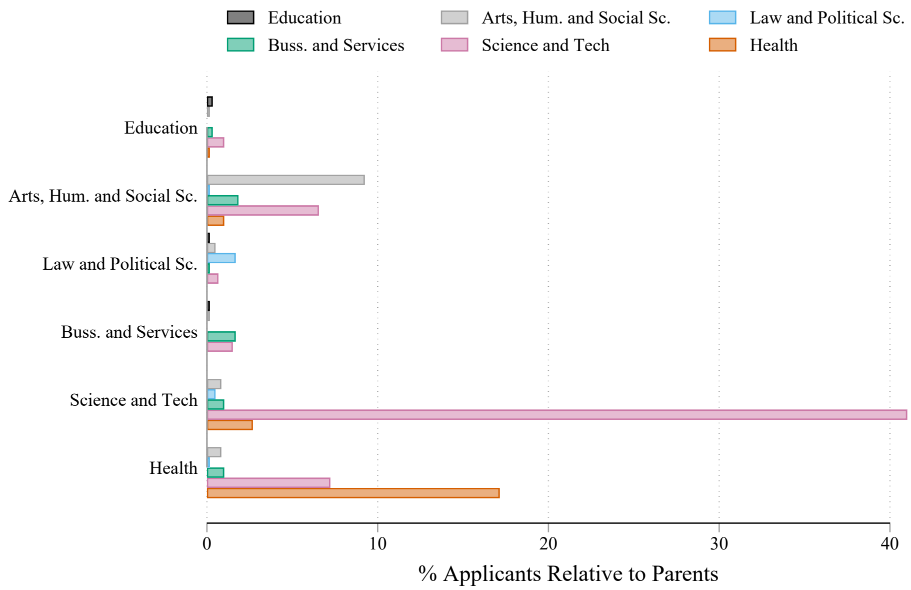
In this section, I analyze the results from the baseline taken in 2019, which provides a nice setting for a descriptive analysis of the program. Every year, MacTec organizes a round of applications for the program starting on January 1st. In 2019, the researchers participated on this process and they modified the application questionnaire, collecting data on 754 applicants. This baseline was mainly designed to capture basic information about families and some indicators of beliefs, stereotypes and aspirations.

I find that 78% of people filling out the survey are the applicant's mother and the average age of parent or tutor is 40 years old. Also, I find that parents are highly educated: 90% of them have done some higher education and what is more, 50% have completed a college degree, which is consistent with the analysis on selection. Only 1% of parents/tutors have a native language different than Spanish (i.e., Quechua or Aymara). Additionally, 30% of applicants live in 4 districts out the 44 in Metropolitan Lima. They come from households with an average size of 4.8 people, so 47% of applicants have one sibling and 25% are single children. Most applicants are in 3rd and 4th grade.

Regarding aspirations and as expected, most parents say they expect their daughters to finish college and then work. I also ask, conditional on going to college, what major they wish their daughter follow and 58% expect their daughters to follow a STEM major. This is also highly correlated with their daughter's expectations, as seen in Figure A.1, where I plot the shares relative to their parents' preferences.

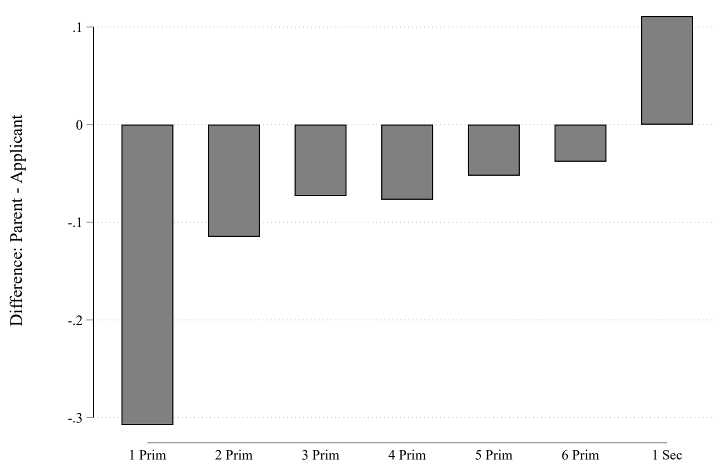
Regarding school, most parents agree their daughters are better at math and science, but 30% of them said that their daughter are good at *all* subjects. When I asked parents how much effort does the applicants put into school, 56% replied *what is needed* while 40% replied *a lot* and the rest, *little or almost none*. However, when comparing to the applicant's responses, (i.e, difference of tutor minus applicant's answers), I find that parents tend to underestimate their daughters' effort. I also find that it corrects overtime: parents underestimate their

Figure A.1: Major Choice Expectations



daughters when they are very young but overestimate when older as seen in Figure A.2. [12]

Figure A.2: Biases on Effort at School by Grade

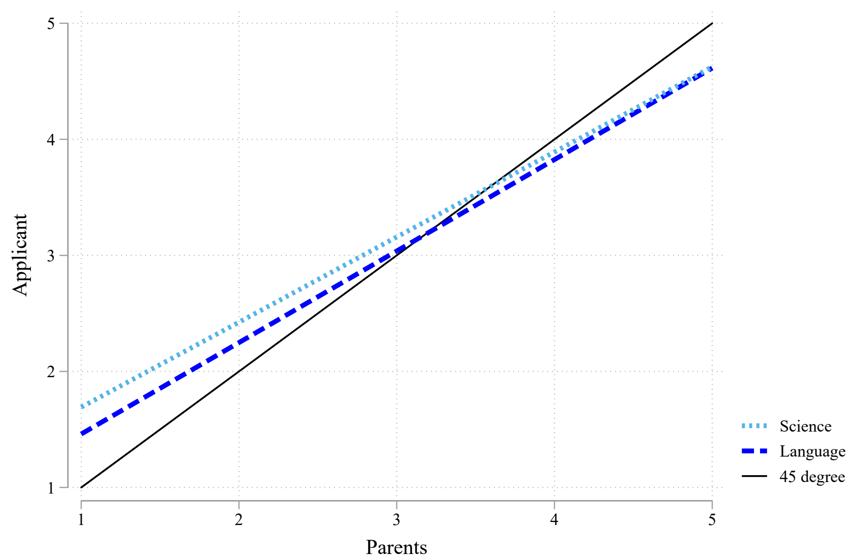


In terms of academic achievement, we find that most parents think their daughter gets better grades in STEM subjects and 30% said that the applicant is good in *all* subjects.

[12]Read Figure A.2 carefully, both 1st graders of primary and secondary school have very little observations so I cannot extract meaningful conclusions from both extremes.

This will suggest that parents tend to overestimate their daughters in both academic performance and ability. When comparing with the applicant's answers - where we ask the same questions regarding ability in some subjects (STEM vs non-STEM), we find some biases. As Figure A.3 shows, girls tend to be more optimistic relative to their parents for lower levels of ability while in higher levels, girls tend to be more pessimistic relative to their parents. And what is more, there is a small sign of overestimating abilities in non-STEM and underestimating on STEM (parents relative to applicants). Both measures are significant and positively correlated, meaning that parents over/under estimate their daughters ability both in STEM/non-STEM together.

Figure A.3: Biases on Ability - Applicants relative to Parents



I also ask about stereotypes, mainly if there are biases for against women in STEM and non-STEM subjects, and I do not find significant differences. This might be because the questions were too straight forward, and the ”right” answer is quite obvious. I ask the same question both applicants and their parents/tutors, so I compared their answers relative to each other. In both questions, most of parents and applicants said that both women and man are equally good, but this was more intense among parents. There was 20% of applicants said girls were better suggesting they might be more positively biased than their parents. Finally, I asked applicants questions about their preferences and aspirations. As expected, most girls reveled that their favorite subject at school were STEM and 30% said *all subjects* were their favorite.

Figure A.4: Major Choice Preference

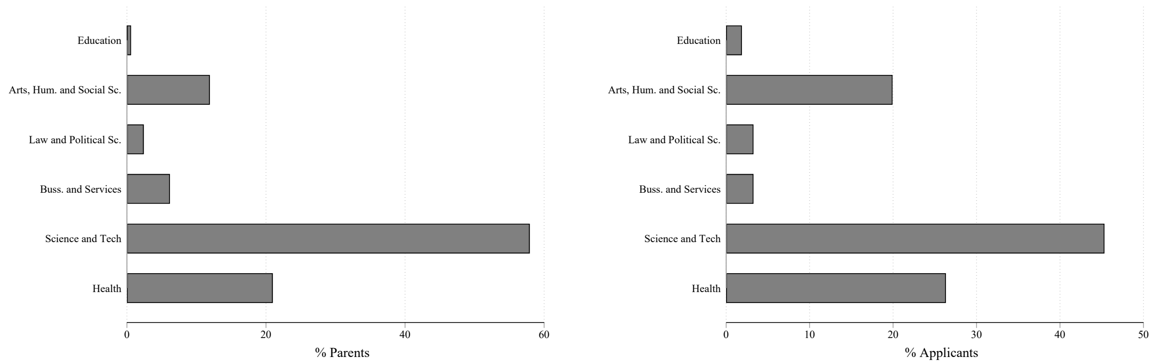


Table A.1: Applicants' Favorite Subject

Subject	Count	Percent
Math	248	33
Science	222	29
Language	61	8
Social Sciences	66	9
All	104	14
None	8	1
Other	45	6
Total	754	100

Table A.2: For which subject does your daughter have better academic achievement?

Subject	Number	Percent
Math	246	33
Science	181	24
Language	62	8
Social Sciences	32	4
All	224	30
None	2	0
Other	7	1
Total	754	100

Figure A.5: Biases on Effort at School

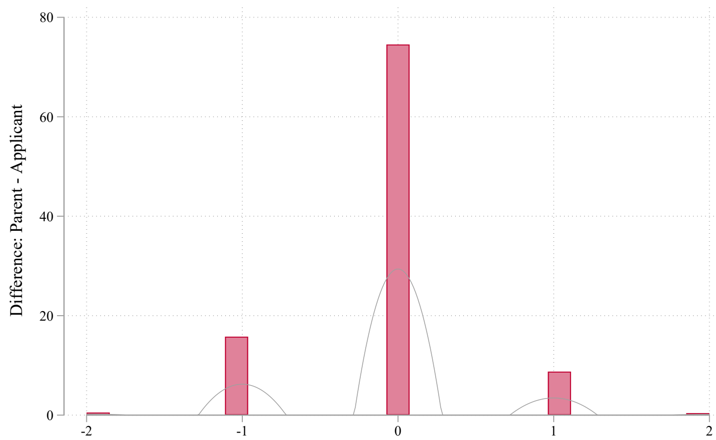


Figure A.6: Beliefs on Ability - Applicants

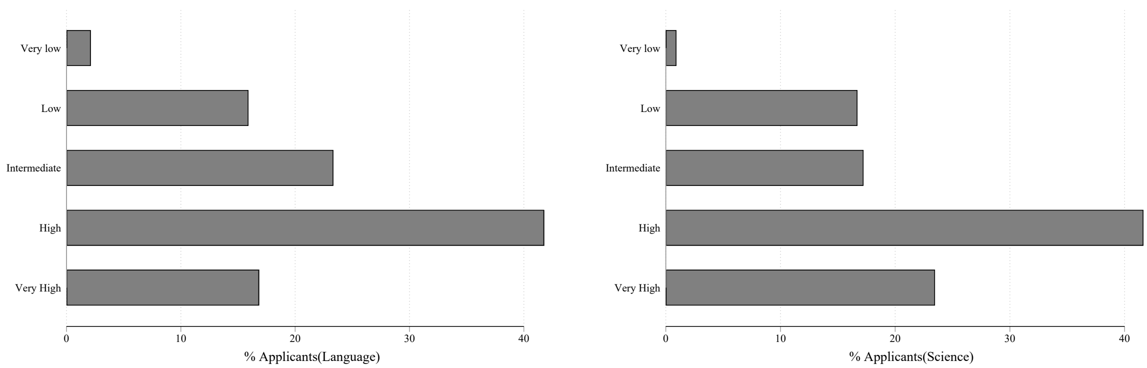


Figure A.7: Beliefs on Ability - Tutors

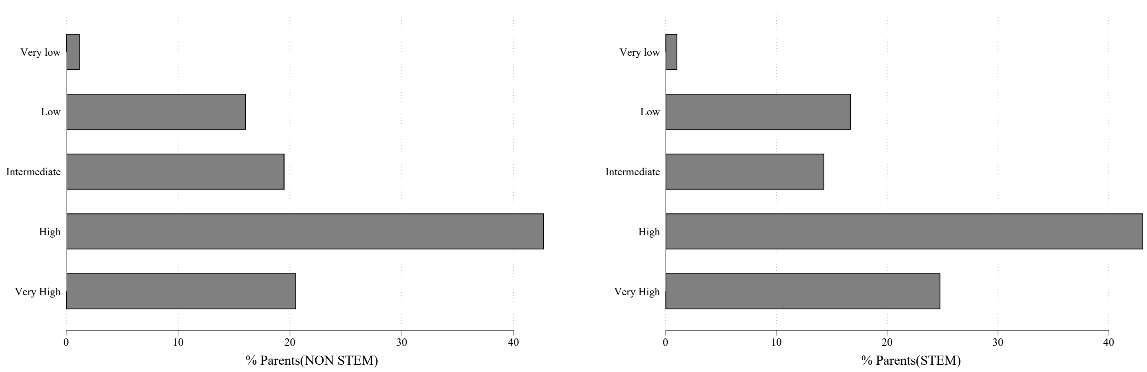


Table A.3: Who does it better?- Applicant

Statement	No STEM Count	Percent	STEM Count	Percent
Girls are better	211	28	216	29
Boys are better	34	5	29	4
All are good	475	63	487	65
Dont know	34	5	22	3
Total	754	100	754	100

Table A.4: Who does it better?- Parents/Tutor

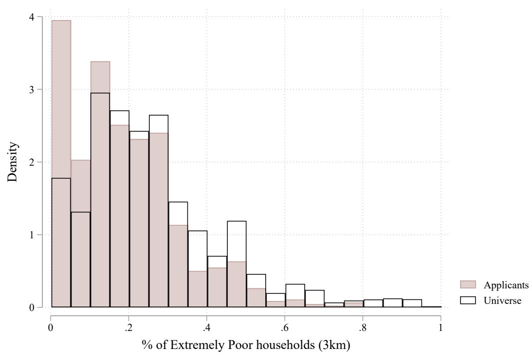
Statement	No STEM Count	Percent	STEM Count	Percent
Women are better	69	9	50	7
Men are better	6	1r	20	3
All are good	674	89	675	90
Dont know	5	1	9	1
Total	754	100	754	100

A.2 Additional Tables and Figures

A.2.1 Selection

Figure A.8: Selection on School Characteristics as a Socio-Economical Measure

(a) Extreme Poverty Share



(b) Private School Fees

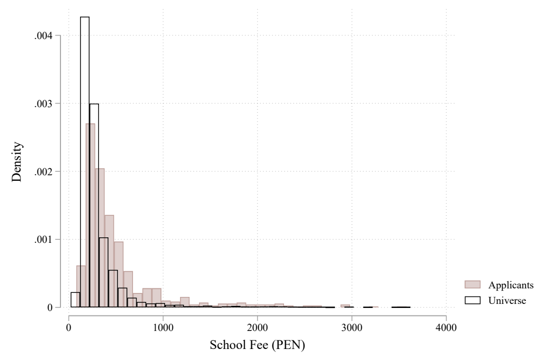


Figure A.9: Where do the applicants come from?

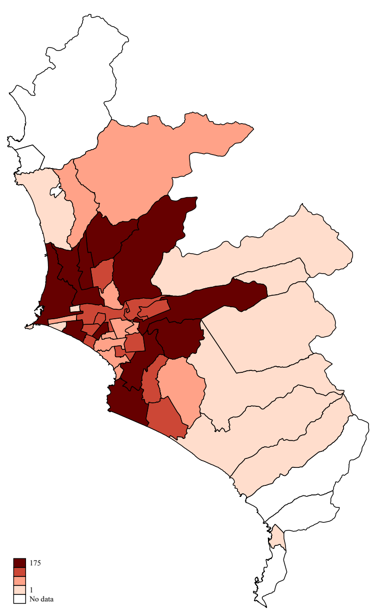
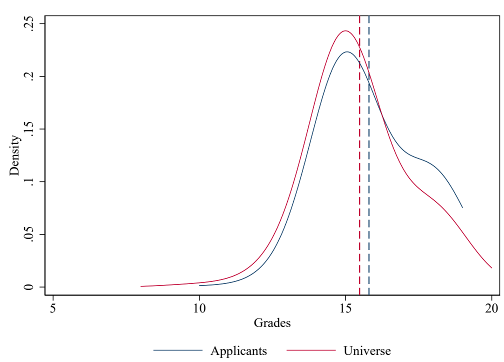
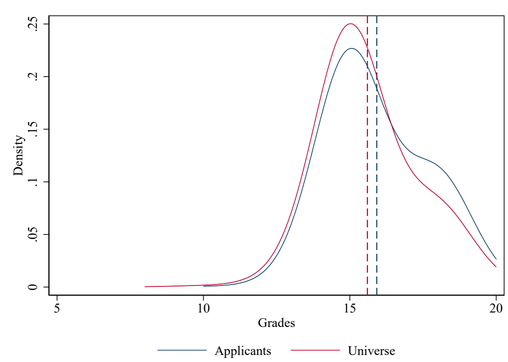


Figure A.10: School Grades Selection in Non-STEM Subjects

(a) Spanish



(b) Social Sciences



A.2.2 Randomization

Figure A.11: Geographical Distribution of Applicants

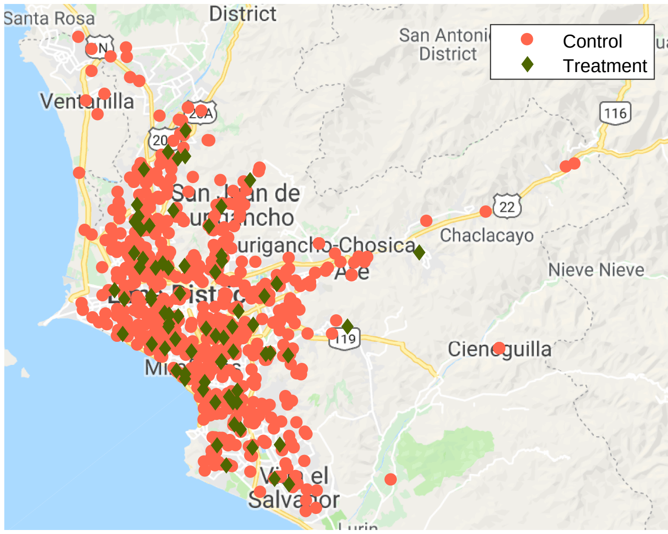
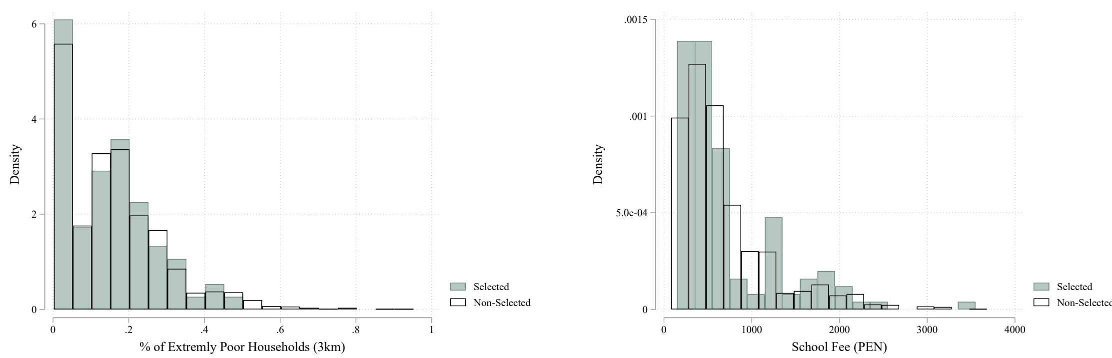


Figure A.12: School as a Socio-Economical Measure

(a) Extreme Poverty Around School

(b) Private School Fees



A.2.3 Endline Implmentation

Table A.5: Balance Control and Treatment on Endline Survey

Variable	(1) Control	(2) Treatment	(3) Difference
Applicant's age	9.612	9.464	-0.092
	(1.187)	(1.039)	(0.137)
Applicant's grade	4.547	4.414	-0.135
	(1.354)	(1.247)	(0.160)
Parent/tutor completed college or more	0.618	0.598	-0.023
	(0.487)	(0.492)	(0.059)
School is public	0.161	0.211	0.060
	(0.368)	(0.410)	(0.047)
Poverty Around School	0.158	0.143	-0.010
	(0.129)	(0.116)	(0.016)
School fee (PEN)	606.5	625.5	1.390
	(479.2)	(467.4)	(67.12)
ECE Spanish	645.9	645.9	-0.262
	(43.19)	(39.78)	(5.314)
ECE Math	610.0	610.3	1.491
	(57.02)	(52.75)	(7.026)
No. Teachers	23.90	23.9	-1.077
	(16.69)	(17.38)	(2.054)
No. Students	424.6	421.2	-14.33
	(295.1)	(335.3)	(37.68)
Observations	238	112	350

A.2.4 Academic Achievement Results

Table A.6: Academic outcomes during treatment year

	(1) Arts	(2) Spanish	(3) Religion	(4) Math	(5) Science	(6) Social Sciences
Treated	-0.153	-0.226*	-0.186	-0.253*	-0.004	-0.014
	(0.135)	(0.133)	(0.132)	(0.131)	(0.136)	(0.133)
Mean	16.26	15.96	16.07	15.90	15.96	15.99
Obs.	1549	1549	1518	1549	1549	1546

Including year FE. Significance: *** p<0.01 ** p<0.05 * p<0.1.

Table A.7: Academic outcomes one year after treatment

	(1) Arts	(2) Spanish	(3) Religion	(4) Math	(5) Science	(6) Social Sciences
Treated	-0.046	-0.259	-0.001	-0.166	-0.220	-0.340**
	(0.190)	(0.200)	(0.190)	(0.173)	(0.157)	(0.172)
Mean	16.26	15.96	16.07	15.90	15.96	15.99
Obs.	1143	1144	1076	1144	1144	1137

Including year FE. Significance: *** p<0.01 ** p<0.05 * p<0.1.

Table A.8: Academic outcomes two years after treatment

	(1) Arts	(2) Spanish	(3) Religion	(4) Math	(5) Science	(6) Social Sciences
Treated	-0.244	-0.373*	-0.192	0.183	-0.251	-0.264
	(0.238)	(0.221)	(0.205)	(0.225)	(0.228)	(0.179)
Mean	16.26	15.96	16.07	15.90	15.96	15.99
Obs.	519	519	496	519	519	513

Including year FE. Significance: *** p<0.01 ** p<0.05 * p<0.1.

Figure A.13: Relative Grades Results: During the treatment year

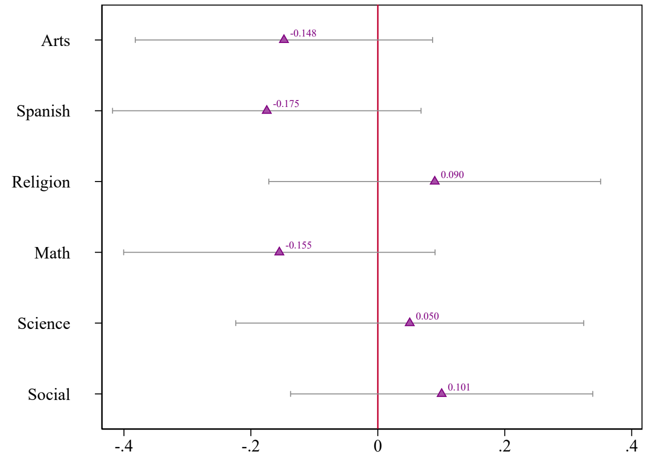


Figure A.14: Relative Grades Results: Including 1 year after treatment

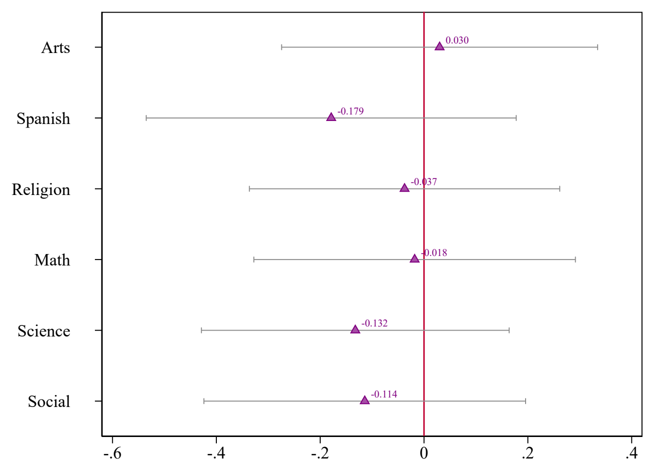
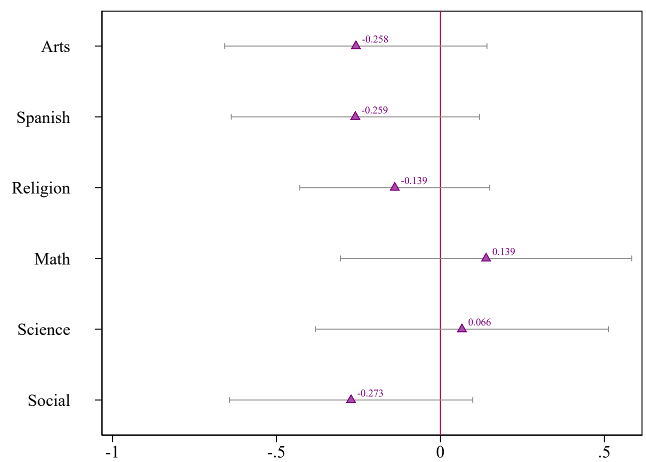


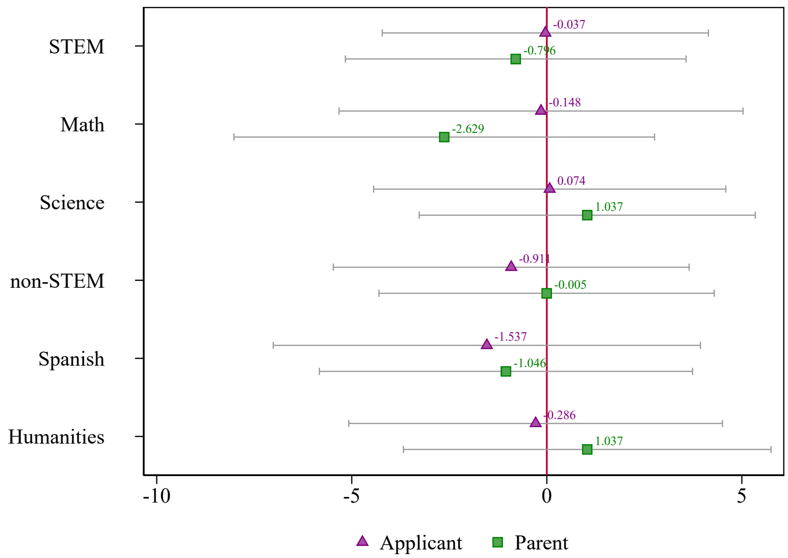
Figure A.15: Relative Grades Results: Including 2 years after treatment



A.2.5 Beliefs about ability, grades and effort at School

Figure A.16: Perceptions on Ability and Grades by Subject

(a) Ability



(b) Grades

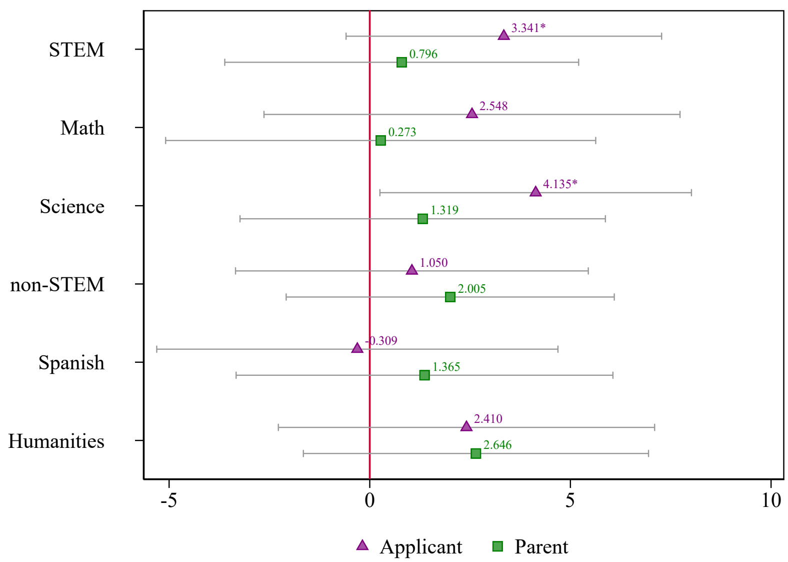


Table A.9: Applicant's self-perception on ability

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	0.204	-0.037	0.151	-0.148	0.257	0.074	-0.735	-0.911	-1.327	-1.537	-0.143	-0.286
	(2.076)	(2.125)	(2.605)	(2.633)	(2.217)	(2.294)	(2.243)	(2.319)	(2.696)	(2.784)	(2.371)	(2.435)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	82.35	82.35	80.14	80.14	84.55	84.55	81.27	81.27	81.57	81.57	80.97	80.97
Obs.	338	337	338	337	338	337	338	337	338	337	338	337

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.10: Parent's perception on daughter's ability

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	-0.515	-0.796	-2.076	-2.629	1.047	1.037	-0.035	-0.005	-0.868	-1.046	0.798	1.037
	(2.197)	(2.220)	(2.711)	(2.741)	(2.138)	(2.191)	(2.139)	(2.185)	(2.403)	(2.431)	(2.331)	(2.396)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	84.65	84.65	82.73	82.73	86.58	86.58	80.69	80.69	81.42	81.42	79.97	79.97
Obs.	350	349	350	349	350	349	350	349	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.11: Parent-applicant difference on ability perception

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	-0.937	-1.026	-2.248	-2.559	0.374	0.508	0.588	0.709	0.546	0.450	0.631	0.967
	(1.974)	(1.916)	(2.317)	(2.250)	(2.281)	(2.267)	(2.038)	(2.027)	(2.641)	(2.696)	(2.282)	(2.253)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	2.421	2.421	2.655	2.655	2.188	2.188	-0.419	-0.419	0.0218	0.0218	-0.860	-0.860
Obs.	338	337	338	337	338	337	338	337	338	337	338	337

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.12: Applicant's self-perception on grades

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	3.509*	3.341*	2.924	2.548	4.094**	4.135**	1.070	1.050	-0.144	-0.309	2.284	2.410
	(1.975)	(1.999)	(2.600)	(2.635)	(1.923)	(1.973)	(2.196)	(2.234)	(2.499)	(2.541)	(2.319)	(2.383)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	81.19	81.19	79.44	79.44	82.94	82.94	80.28	80.28	81.21	81.21	79.35	79.35
Obs.	338	337	338	337	338	337	338	337	338	337	338	337

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.13: Parent's perception on daughter's grades

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	1.112	0.796	0.672	0.273	1.551	1.319	2.003	2.005	1.598	1.365	2.408	2.646
	(2.178)	(2.242)	(2.653)	(2.724)	(2.245)	(2.314)	(2.001)	(2.078)	(2.308)	(2.387)	(2.100)	(2.186)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	83.62	83.62	81.88	81.88	85.36	85.36	80.33	80.33	81.25	81.25	79.42	79.42
Obs.	350	349	350	349	350	349	350	349	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.14: Parent - applicant difference on grades perception

	STEM						Non-STEM					
	Avg.		Math		Science		Avg.		Spanish		Humanities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	-2.443	-2.655*	-1.942	-2.062	-2.944	-3.248*	0.808	0.787	1.658	1.532	-0.041	0.042
	(1.507)	(1.514)	(1.703)	(1.675)	(1.917)	(1.952)	(1.757)	(1.788)	(2.242)	(2.284)	(1.855)	(1.886)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	2.493	2.493	2.389	2.389	2.598	2.598	0.0808	0.0808	0.0873	0.0873	0.0742	0.0742
Obs.	338	337	338	337	338	337	338	337	338	337	338	337

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** $p<0.01$ ** $p<0.05$ * $p<0.1$

Figure A.17: Overconfidence in Spanish

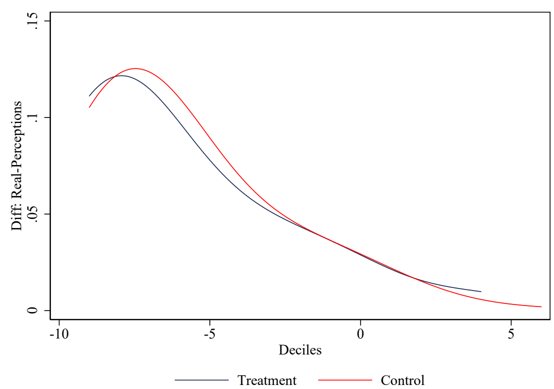


Figure A.18: Overconfidence Humanities and Social Sciences

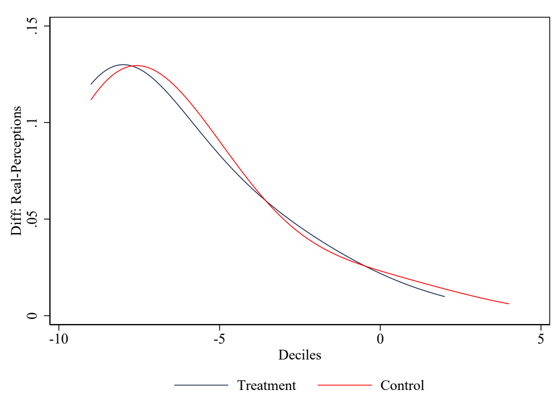
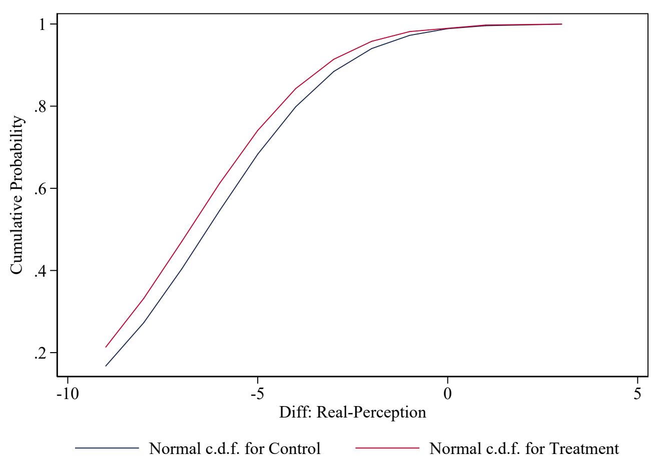


Figure A.19: Overconfidence in Science and CDF by Treatment Status



A.2.6 Time Use

Table A.15: Time Use Effects, no controls

	Sleeping	Study and School Homework	Eating	Personal Hygiene	Social Media	Play or Watch TV	Educational Shows	Books or Internet	Work/Help Parents	Household Chores	Personal Projects	Sport/Arts Training	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Treated	-0.329	-0.401*	-0.038	0.058	-0.127	0.126	0.023	0.014	0.071	0.081	0.227**	0.070	0.225*
	(0.273)	(0.229)	(0.104)	(0.107)	(0.185)	(0.137)	(0.129)	(0.111)	(0.097)	(0.080)	(0.103)	(0.101)	(0.127)
Controls	No	No	No	No	No	No	No	No	No	No	No	No	No
Mean	7.748	5.351	1.761	1.203	1.180	1.212	1.077	1.113	0.595	0.869	0.698	0.910	0.284
Obs.	327	327	327	327	327	327	327	327	327	327	327	327	327

All models have year FE.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table A.16: Time Use Effects, with controls

	Sleeping	Study and School Homework	Eating	Personal Hygiene	Social Media	Play or Watch TV	Educational Shows	Books or Internet	Work/Help Parents	Household Chores	Personal Projects	Sport/Arts Training	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Treated	-0.373	-0.399*	-0.028	0.075	-0.092	0.110	0.014	0.033	0.090	0.074	0.221**	0.066	0.210
	(0.271)	(0.229)	(0.104)	(0.110)	(0.181)	(0.140)	(0.127)	(0.110)	(0.098)	(0.079)	(0.105)	(0.103)	(0.127)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	7.748	5.351	1.761	1.203	1.180	1.212	1.077	1.113	0.595	0.869	0.698	0.910	0.284
Obs.	326	326	326	326	326	326	326	326	326	326	326	326	326

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

B Perceptions and Major Choice

B.1 Perceptions in STEM Majors

Table B.1: Applicant's perception of STEM major

	STEM avg.		Eng.		Medicine		Maths		Architecture	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	-6.592**	-6.462**	-5.297	-4.819	-4.000	-3.816	-8.075**	-7.914**	-8.996**	-9.300**
	(2.710)	(2.771)	(3.846)	(3.919)	(4.122)	(4.178)	(3.927)	(3.997)	(3.929)	(4.008)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	65.50	65.50	67.67	67.67	65.78	65.78	61.61	61.61	66.91	66.91
Obs.	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.2: Parent's perception of STEM major

	STEM avg.		Eng.		Medicine		Math		Architecture			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
Treated	0.001	0.015	-0.353	-0.622	0.607	0.753	-0.970	-0.830	-2.334	-3.103	1.285	0.690
	(0.060)	(0.061)	(2.449)	(2.485)	(2.819)	(2.894)	(3.286)	(3.313)	(3.545)	(3.607)	(3.355)	(3.397)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	0.513	0.513	78.82	78.82	85.80	85.80	81.95	81.95	71.11	71.11	76.42	76.42
Obs.	350	349	350	349	350	349	350	349	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.3: Parent-applicant difference perception of STEM major

	STEM avg.		Eng.		Medicine		Math		Architecture	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	6.035**	5.688**	5.697	5.379	3.400	3.393	5.498	4.698	9.546**	9.281**
	(2.726)	(2.750)	(3.698)	(3.796)	(3.756)	(3.774)	(4.410)	(4.429)	(3.798)	(3.874)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	13.10	13.10	18.08	18.08	15.58	15.58	9.202	9.202	9.534	9.534
Obs.	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.4: Applicant's perception about peers in STEM major

	STEM avg.		Eng.		Medicine		Math		Architecture	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	-8.092***	-7.964***	-7.008*	-6.360*	-2.178	-2.215	-10.810***	-10.860***	-12.373***	-12.422***
	(2.928)	(2.995)	(3.710)	(3.770)	(3.805)	(3.851)	(3.913)	(3.921)	(3.628)	(3.702)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	64.08	64.08	64.65	64.65	67.34	67.34	58.07	58.07	66.26	66.26
Obs.	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

B.2 Perception in Non-STEM Majors

Table B.5: Applicant's perception of non-STEM major

	Non-STEM avg.		Law		Business		Journalism		Education		Sociology		Arts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treated	-7.938***	-7.980***	-13.069***	-13.311***	-7.890**	-7.749**	-10.006***	-10.001**	-5.972*	-6.222*	-7.330**	-6.710*	-3.359	-3.887
	(2.580)	(2.617)	(3.823)	(3.808)	(3.874)	(3.903)	(3.803)	(3.863)	(3.605)	(3.699)	(3.532)	(3.591)	(3.631)	(3.679)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	57.39	57.39	55.35	55.35	60.69	60.69	50.86	50.86	49.35	49.35	48.97	48.97	79.14	79.14
Obs.	330	329	330	329	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.6: Parent's perception of non-STEM major

	Non-STEM avg.		Law		Business		Journalism		Education		Sociology		Arts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treated	0.977	0.909	-0.007	0.353	-2.924	-3.022	-0.243	0.096	1.747	1.492	2.795	2.317	4.491	4.219
	(2.828)	(2.866)	(3.784)	(3.806)	(3.430)	(3.463)	(3.874)	(3.894)	(3.725)	(3.756)	(3.702)	(3.730)	(3.370)	(3.438)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	63.68	63.68	64.05	64.05	74.68	74.68	54.83	54.83	58.93	58.93	56.92	56.92	72.63	72.63
Obs.	350	349	350	349	350	349	350	349	350	349	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.7: Parent-applicant difference in perception of non-STEM majors

	Non-STEM avg.		Law		Business		Journalism		Education		Sociology		Arts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treated	9.110***	9.148***	12.797***	13.430***	5.404	5.144	9.722**	10.022**	8.116*	8.269**	10.413***	9.486**	8.213**	8.540**
	(2.870)	(2.853)	(4.410)	(4.306)	(4.003)	(3.963)	(4.090)	(4.200)	(4.156)	(4.117)	(3.966)	(3.926)	(4.113)	(4.224)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	5.442	5.442	8.099	8.099	13.39	13.39	3.502	3.502	8.610	8.610	6.623	6.623	-7.578	-7.578
Obs.	330	329	330	329	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.8: Applicant's perception about peers in non-STEM careers

	Non-STEM avg.		Law		Business		Journalism		Education		Sociology		Arts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Treated	-9.469***	-9.881***	-8.738**	-8.937**	-4.624	-4.609	-7.965**	-9.037**	-13.591***	-14.726***	-10.352***	-10.152***	-8.702**	-8.800**
	(2.563)	(2.547)	(3.819)	(3.816)	(3.692)	(3.711)	(3.652)	(3.657)	(3.583)	(3.508)	(3.690)	(3.717)	(3.696)	(3.685)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	60.53	60.53	57.96	57.96	63.01	63.01	56.65	56.65	56.58	56.58	53.34	53.34	75.62	75.62
Obs.	330	329	330	329	330	329	330	329	330	329	311	310	311	310

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

B.3 Major Choice

Table B.9: Applicant's major choice in STEM

	STEM avg.		Science		Health	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	-0.002	-0.003	0.043	0.038	-0.062	-0.054
	(0.063)	(0.064)	(0.061)	(0.061)	(0.044)	(0.045)
Controls	No	Yes	No	Yes	No	Yes
Mean	0.574	0.574	0.354	0.354	0.202	0.202
Obs.	315	314	330	329	330	329

All models have year FE. Controls include parents' education, applicant's age, and a dummy that indicates whether the girl applied more than once. . Significance: *** p<0.01 ** p<0.05 * p<0.1.

Table B.10: Applicant's major choice in non-STEM

	Non-STEM avg.		Humanities		Law		Education		Business		Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	0.002	0.003	0.010	0.015	-0.037	-0.041*	0.001	-0.002	-0.012	-0.006	0.025	0.025
	(0.063)	(0.064)	(0.053)	(0.054)	(0.023)	(0.023)	(0.013)	(0.013)	(0.027)	(0.028)	(0.030)	(0.030)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	0.426	0.426	0.211	0.211	0.0628	0.0628	0.0135	0.0135	0.0717	0.0717	0.0538	0.0538
Obs.	315	314	330	329	330	329	330	329	330	329	330	329

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.11: Parent's preferred major choice in STEM

	STEM avg.		Science			Health
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.038	0.021	0.020	0.006	0.028	0.027
	(0.056)	(0.056)	(0.060)	(0.060)	(0.043)	(0.044)
Controls	No	Yes	No	Yes	No	Yes
Mean	0.717	0.717	0.513	0.513	0.168	0.168
Obs.	334	333	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

Table B.12: Parent's preferred major choice in STEM

	Non-STEM avg.		Humanities		Law		Education		Business		Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated	-0.038	-0.021	0.019	0.020	-0.005	-0.005	-0.014	-0.012	-0.037	-0.027	0.005	0.007
	(0.056)	(0.056)	(0.040)	(0.040)	(0.016)	(0.015)	(0.010)	(0.010)	(0.035)	(0.035)	(0.018)	(0.018)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean	0.283	0.283	0.0966	0.0966	0.0294	0.0294	0.00840	0.00840	0.113	0.113	0.0210	0.0210
Obs.	334	333	350	349	350	349	350	349	350	349	350	349

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1

B.4 Occupation Choice

Table B.13: Preferred occupation, no controls

	Hard science (1)	Eng. (2)	Health (3)	Econ. (4)	Business Admin. (5)	Business Woman (6)	Law (7)	Humanities (8)	Teacher (9)	Languages (10)	Media (11)	Sports (12)	Military or Police (13)	Chef (14)	Arts (15)	Manual services (16)	Unsure (17)
Treatment	0.057	0.007	-0.056	-0.002	-0.020	0.012	-0.020	0.025	0.012	-0.014*	-0.011	-0.023**	0.008	-0.020	0.100*	0.026	0.038
	(0.049)	(0.056)	(0.056)	(0.011)	(0.022)	(0.022)	(0.026)	(0.022)	(0.019)	(0.008)	(0.026)	(0.009)	(0.012)	(0.027)	(0.054)	(0.018)	(0.025)
Controls	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Mean	0.186	0.245	0.327	0.0136	0.0409	0.0318	0.0682	0.0136	0.0227	0.0136	0.0409	0.0273	0.00455	0.0636	0.182	0.00455	0.0227
Obs.	325	325	325	325	325	325	325	325	325	325	325	325	325	325	325	325	325

All models have year FE. Significance: *** p<0.01 ** p<0.05 * p<0.1.

Table B.14: Preferred occupation, controls

	Hard science (1)	Eng. (2)	Health (3)	Econ. (4)	Business Admin. (5)	Business Woman (6)	Law (7)	Humanities (8)	Teacher (9)	Languages (10)	Media (11)	Sports (12)	Military or Police (13)	Chef (14)	Arts (15)	Manual services (16)	Unsure (17)
Treated	0.058	-0.000	-0.055	-0.001	-0.012	0.014	-0.020	0.029	0.010	-0.014*	-0.009	-0.021**	0.007	-0.027	0.100*	0.029	0.043*
	(0.048)	(0.056)	(0.058)	(0.012)	(0.021)	(0.024)	(0.027)	(0.023)	(0.019)	(0.009)	(0.026)	(0.009)	(0.011)	(0.028)	(0.054)	(0.019)	(0.025)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.186	0.245	0.327	0.0136	0.0409	0.0318	0.0682	0.0136	0.0227	0.0136	0.0409	0.0273	0.00455	0.0636	0.182	0.00455	0.0227
Obs.	324	324	324	324	324	324	324	324	324	324	324	324	324	324	324	324	324

All models have year FE.
Controls: parents' education, applicant's age, and a dummy if the girl applied more than once.
Significance: *** p<0.01 ** p<0.05 * p<0.1