Can New Learning Opportunities Reshape Gender Attitudes for Girls?: Field Evidence from Tanzania*

So Yoon Ahn — University of Illinois at Chicago

Youjin Hahn — Yonsei University

Semee Yoon — Yonsei University

August 26, 2021

Abstract

We study how educational opportunities change adolescents' gender attitudes in Tanzania, using an experiential education program focused on STEM subjects. After the intervention, girls' gender attitudes became more progressive by 0.29 standard deviations, but boys' gender attitudes did not change. Perceived improvement in the labor market opportunities appears to be an important channel to explain the result. The intervention also increased girls' weekly study hours and boosted their interests in STEM-related subjects and occupations. Our results show that providing STEM-related educational opportunities to girls in developing countries can be an effective way of improving their gender attitudes.

JEL classifications: I25, J13, J16

*We are grateful for Ji-Young Rhee, Irene Summary, Mary Paschal, Joshua Dennis, and Robert Kundi from *e3empower*, without whom our study would not have been possible. We thank Ben Feigenberg, Yiqun Chen, Suanna Oh, Ben Ost, Javaeria Qureshi, Steve Rivkin, Hee-Seung Yang, Jiso Yoon, and various seminar participants for helpful comments, and Eliud Moses and Janeth Shayo for providing support from Compassion. Yujin Chung, Mei Meiqing, and Ann James provided excellent research assistance. This trial received an IRB approval from Yonsei University (Approval number: 7001988-201812-HR-515-03) and the pre-analysis plan is registered in the AEA RCT Registry (AEARCTR-0005215). This research was supported by the Yonsei Signature Research Cluster Program of 2021 (2021-22-0011) and the 2018 Social Engagement Fund, provided by the Institute of Global Engagement and Empowerment and at Yonsei University. The data and codes will be available for replication upon publication.

1 Introduction

Gender inequality is widespread in developing countries in every aspect of life. Women typically complete fewer years of schooling, earn less, and lack decision-making power compared to men. (UN Statistics Division [2018]). For example, in Tanzania, where our study was conducted, enrollment rates in upper secondary education are approximately 50% higher for boys than girls (Tanzania National Bureau of Statistics [2018]).[1] One of the contributing factors for this long-lasting gender inequality is deep-rooted gender norms that favor men (Jayachandran [2015]; Jayachandran [2020]).[2] Therefore, reshaping gender attitudes serves as an important step towards achieving gender equality in various dimensions. Moreover, studies find a strong relationship between the gender attitudes of mothers and those of their daughters, suggesting an inter-generational transmission of gender attitudes whereby shifts in attitude in the present have positive long-term and inter-generational implications (Moen et al. [1997]; Dhar et al. [2019]).

In particular, existing studies have shown that education programs directed towards gender equality, such as school-based gender equality education (Dhar et al. [2020]) and life skill interventions (Edmonds et al. [2020]), effectively induce more gender-equitable views. Gender attitudes are also affected by factors that are not directly designed to change gender norms; studies show that labor market opportunities (Jensen [2012]), TV programming (Jensen and Oster [2009]; La Ferrara et al. [2012]), and female quota in political institutions (Beaman et al. [2009]) affect gender attitudes and fertility preferences.

In this paper, we study whether new learning opportunities that focus on hands-on experiences reshape gender attitudes among Standard 7 to Form 3 (grade 7 to grade 10) students

[1] In Tanzania, the median gender pay gap in hourly wages is 25% (International Labour Office [2018]). Moreover, widows and daughters are not guaranteed inheritance rights (UN Statistics Division [2018]).

[2] Gender norms that favor men are not limited to developing countries. Bertrand et al. [2015] show there exists a social norm that wives should not earn more than her husbands in the United States. Bertrand et al. [2020] discuss negative social attitudes toward working women in developed countries. Alesina et al. [2013] study historical origins of skewed gender norms by examining children of immigrants living in Europe and the U.S. Fewer girls are born among U.S. residents of East Asian (Almond and Edlund [2008]) and South Asian legacy (Abrevaya [2009]).

in the Tabora region of Tanzania.[3] Specifically, we analyze the impacts of providing a Science, Technology, Engineering, and Mathematics (STEM) education program as a 5-day intensive boot camp during the school break in mid-November of 2018. The purpose of the STEM boot camp is to provide an experiential learning opportunity to youth to hone their creativity, computational thinking, and problem-solving skills. This program is specifically designed to provide hands-on experiences in STEM such as computer coding, building robots, and creating simple apps. We examine whether the STEM boot camp influenced non-cognitive outcomes, including gender attitudes, self-esteem, and educational and long-term career aspirations in STEM, beyond generating knowledge of coding and robot technology. In this study, we target students in their early adolescence, because the literature suggests that interventions during adolescence can be particularly effective at changing attitudes[4]; during this period, youth form their own value systems, but they still have flexible attitudes (Crone and Dahl [2012]).

While our intervention does not target gender attitudes, it may affect gender norms through several channels. For girls, as they tend to find STEM subjects more difficult than boys do (Semali and Mehta [2012]), our intervention focusing on more accessible hands-on experiences on STEM may boost girls' self-esteem and their interests in relevant subjects (Rogers and Portsmore [2004]; Nourbakhsh et al. [2005]). This may lead to changes in their attitudes towards further education opportunities and women's social participation. The STEM intervention may also change labor market aspirations and perceived earnings if skills imperative to STEM are deemed to generate a higher return in the labor market. Given that STEM field is typically male-dominated (e.g. Kahn and Ginther [2017]), we expect that these channels would particularly affect girls' gender attitudes. However, several features of our intervention may affect both boys' and girls' gender attitudes. For instance, some students in the treatment group were exposed to a female instructor, and by being exposed to female role models, attitudes of boys and girls may change (Del Carpio and Guadalupe [2018]; Kipchumba et al.

[3] In Tanzanian education system, Standard 1 to Standard 7 (grade 1 to grade 7) students are in primary schools and Form 1 to Form 4 (grade 8 to grade 11) students are in secondary schools. The education system in Tanzania is depicted in Table A1.

[4] See Koch et al. [2015] for review.

[2021]). Moreover, our intervention involves a mixed-gender environment, and interaction with students of the opposite sex while solving STEM problems may reshape gender attitudes for boys and girls.

Although a positive correlation between education and gender attitudes is widely documented (Kane [1995]; Leaper et al. [2012]; Auletto et al. [2017]), evidence regarding whether educational opportunities lead to more progressive gender attitudes is scarce. It has been difficult to demonstrate a causal impact because education is associated with many other attributes, such as socioeconomic status, family structures, and mothers' gender attitudes, which may also affect gender attitudes. Moreover, it is even more challenging to find contemporaneous measures of gender attitudes that align with educational shocks. Recently, Du et al. [2020] showed that a Chinese compulsory education reform that increased the years of education induced more egalitarian gender role attitudes. They measured the impacts of the educational reform after more than 20 years. Thus, they estimated the broader impact of education including the indirect impacts through actual labor market participation and differential media use, both of which were affected by the reform. Our paper departs from Du et al. [2020] in two aspects. First, we aim to measure more direct impacts of education; and second, the educational content we provide is related to STEM, the subjects typically associated with gender-related stereotypes (Makarova et al. [2019]).

To overcome these aforementioned challenges, we exploit a unique setting in Tanzania where a STEM education program was offered to a subset of students enrolled in a large child sponsorship organization, Compassion International, in the Tabora Urban district in the Tabora region. We surveyed students between grades 7 and 10 at Tabora Compassion centers, as shown in Figure 1, regardless of the participation in the STEM program, before and after the intervention. In particular, we asked questions designed to measure gender attitudes and created a gender-attitudes index; these questions included attitudes towards equal educational opportunities, labor market participation, and traditional family gender roles. We also collected rich baseline characteristics including pre-treatment gender attitudes, educational goals, self-esteem, personalities, labor market aspirations, and detailed household characteristics. Compassion reported that they selected program participants randomly from a larger pool of registered students from the Tabora Compassion Centers. The randomization was done casually, wherein a human decided the treatment status using their best judgment. The extensive communication with Compassion suggest that the possibility of self-selection is limited. Based on this information, we make a conditional independence assumption (CIA); the treatment is exogenous, conditional on the extensive set of baseline characteristics. With the CIA, we can identify the causal impact of educational opportunities on gender attitudes. Although it is not possible to directly test the CIA, the method outlined in Oster [2019] and Altonji et al. [2005] allows us to assess the degree of selection on unobservables based on the degree of selection on observables. Our analysis indicates that the scope of selection on unobservables is limited once we control the pre-treatment gender attitudes.

We find that STEM education makes girls' gender attitudes more progressive by 0.29 standard deviations, whereas it does not affect boys' gender attitudes. We do not find any evidence that this positive result for girls is driven by their tendency to give socially desirable answers. The change in girls' attitudes is accompanied by a boost in self-esteem, stronger interests in STEM subjects, and increased perceived earnings in the labor market. The decomposition analysis indicates that the perceived improvements in labor market opportunities play the most important role in explaining the positive impact on girls, although other factors also partly explain the results. The larger effect for girls is expected given that STEM remains largely dominated by men and boys' improvements in self-esteem or labor market aspirations are less likely to be related to changes in gender attitudes. Interactions with female teachers or students of the opposite-sex may affect boys' gender attitudes, though we find little support for these. Interestingly, our finding is in contrast to the findings on the impacts of a gender-equality education on gender attitudes; a more direct intervention has a stronger impact on boys (Dhar et al. [2020]).

To understand whether changes in girls' gender attitudes lead to behavioral changes, we examine study hours on different subjects and outcomes related to future choices such as educational aspirations, i.e., expected highest level of education and preferred majors, as well as the desired marriage age and number of children. We find a significant increase in study hours for girls; for treated students, the weekly study hours spent on STEM subjects increased by two hours after the intervention. Regarding future choices, although levels of educational aspirations did not change, girls who participated in the program were more likely to choose engineering or technology as more preferred majors at college compared to the girls in the control group. Moreover, treated girls have a lower propensity to pursue a job that requires relatively low skills, such as agriculture, domestic service and other manual jobs. We did not find any impacts on desired marriage age, but treated girls desired fewer children, consistent with the findings that they expect to earn more and they would have a higher opportunity cost of having children if they secure higher-paying jobs.

This paper contributes to our understanding of gender-related preferences. Existing studies show the effects of gender norms on economic decisions (e.g., Bertrand et al. [2015]; Duflo [2012]; Bertrand et al. [2020]). Gender norms are believed to be persistent, but the literature shows that gender preferences can be changed by various factors, including gender equality interventions[5] and labor market opportunities (e.g., Dhar et al. [2020]; Jensen [2012]). We contribute to the literature by providing evidence that new learning opportunities, especially in areas where girls have limited awareness and knowledge, can significantly improve girls' gender attitudes in developing countries. This highlights another dimension of the benefits associated with providing more educational opportunities for girls.

Moreover, we add to the growing literature by studying the impacts of technology-related education on learners. Most existing papers have focused on the effects of technology on academic outcomes such as school grades and attendance (Malamud and Pop-Eleches [2011]; Fairlie and Robinson [2013]; Bulman and Fairlie [2016]).[6][7] We contribute to the literature by

[5] There is a growing literature on life skill type of interventions to improve girls' outcomes and biased gender attitudes. See Buchmann et al. [2017], Ashraf et al. [2020], Bandiera et al. [2020], Dhar et al. [2020], Edmonds et al. [2020], and McKelway [2020].

[6] One exception is Bauernschuster et al. [2014] who investigated the impacts of Internet on social engagements, showing no evidence that Internet usage affects social capital.

[7] The impacts of technologies on education are theoretically ambiguous because investments in technologies are costly so that they can offset other educational inputs. Also, time inputs in technologies can replace time invested in traditional education (Bulman and Fairlie [2016]). The evidence tends to suggest that with supplementary funding provided, technologies have positive impacts on education (Malamud and Pop-Eleches

showing that technology education has significant impacts on non-cognitive outcomes including gender attitudes, self-esteem, and aspirations in labor market. Studies also show that a robot education program for high schoolers in the U.S. and elementary school curriculum incorporating engineering aspects boosted girls' confidence in technology and interests in STEM (Nourbakhsh et al. [2005]; Rogers and Portsmore [2004]). Given that girls tend to be less confident about STEM subjects, our finding of positive impacts on self-esteem and enhanced interests in STEM subjects shows that benefits associated with these programs extend beyond the first-order outcomes associated with technical education itself.

Our findings have a broader implication on education policy in the Least Developed Countries (LDCs). Technology education can better equip the LDCs' labor forces, regardless of gender, for the digital era. Our results suggest that STEM education interventions do not just provide technical knowledge and skills; these programs further empower girls with more progressive gender views, which may help close pervasive gender gaps. Moreover, the program of our study shows that STEM education programs can be provided at a relatively low cost.[8] The per student cost should decrease even further when these programs are implemented at scale, given the initial fixed costs.

This paper proceeds as follows. Section 2 provides background in Tanzania. Section 3 describes our intervention and data collection. Section 4 presents our empirical challenges and strategies. Section 5 presents the results, and Section 6 concludes.

2 Background in Tanzania

Located on the eastern coast of sub-Saharan Africa and with population greater than 58 million, Tanzania has experienced annual economic growth rates that are above 5% in the past six years, and its GDP per capita was $1,080 in 2019 (World Bank [2021]). While Tanzania has achieved near-universal access to primary education in 2007 (UNICEF [2017]), the completion rate (69%) needs improvement (World Bank [2018]). Moreover, an even lower proportion of students enroll and complete secondary schools, where students are given opportunities to take

[2011]; Fairlie and Robinson [2013]; Bauernschuster et al. [2014]; Dynarski et al. [2007]).

[8] The cost of intervention per student in this study was USD 40.

math and science classes, such as chemistry, biology, and physics (Nuffic [2015]).

There exists a significant gender gap in the education and labor market in Tanzania. On the gender equality index, Tanzania ranked 140th out of 189 countries in 2019 (UNDP [2019]). According to the World Bank [2021], the gross enrollment rates are similar for boys and girls in lower secondary schools, but there are gender disparities in drop-out rates and learning outcomes. For example, approximately 21 percent of boys and only 16 percent of girls that completed lower secondary schooling went onto join upper secondary in 2015-2018. The gender gap persists in the labor market as well; the mean monthly income was 68% higher for men than for women in 2014. The pay gap exists in every type of employment, including paid-employment, self-employment, and agricultural employment (Tanzania National Bureau of Statistics [2017]).

Even though the Ministry of Education, Science, and Technology (MOEST) has strived to provide quality education, making progress has not been easy.[9] As in many other developing countries, Tanzania faces challenges in promoting science education. According to United Nations Educational, Scientific and Cultural Organization (UNESCO)'s needs assessment on the state of science education in Tanzania, primary school science and math teachers are poorly trained and are typically unable to provide real-world examples. Moreover, a significant shortage of adequate laboratories and equipment in most secondary schools makes it difficult to provide quality science education (UNESCO [2009]). For example, there are 0.04 computers per student at secondary schools (The United Republic of Tanzania President's Office [2016]).[10]

The gap in knowledge transfer can also be observed through graduation exam scores. The proportions of students that received pass ratings for their Form 4 (11th grade) Examinations for STEM subjects are lower than those for humanities subjects, such as history, Kiswahili, and English. Moreover, there are significant gender achievement gaps in math and science subjects; in 2018, for math, the pass percentages of male and female students were 24.3% and

[9] For example, the MOEST included the percentage of science and math graduates in higher education as one of the performance indicators.

[10] In the US, the ratio of students to instructional computers with internet access was 3.1 in 2008 (National Center for Education Statistics [2018]).

15.9%, respectively, and for physics, they were 54.4% and 34.9%, respectively (MOEST [2019]). Science subjects are also perceived as difficult and girls are generally more discouraged from taking them (Semali and Mehta [2012]).

Given the lack of quality science education in Tanzania, our STEM education intervention is unique because students are exposed to science education through an immersive experience. Students learn extensive real-world applications with computers, coding programs, and robot education equipment. This departs from the traditional textbook-based education common in Tanzania (Semali and Mehta [2012]). Moreover, girls and boys were equally likely to receive the STEM education program. Details of the intervention are given in the next section.

3 Intervention and data collection

3.1 Setting

The idea to provide high-quality educational opportunities in STEM grew out of a shared sense among three partners of this project: Compassion International, a sponsor organization that aspires to improve livelihoods for youth in low-income settings; e3empower, a local social enterprise with a mission to equip more youth with technical skills; and the authors of this paper, the academic partner interested in the impact of STEM education on non-cognitive outcomes.

As an international non-profit organization that operates in 25 countries world-wide, Compassion International selects children to be sponsored at each local center. The number of children registered has grown to more than 2 million (Compassion [2020]). In Tanzania, Compassion has been working with 506 churches, referred to as local centers, throughout the country to provide resources, skills, and opportunities to more than 100,000 children so that they can reach their full potential (Compassion [2021]). Currently, Compassion selects sponsored children based on needs and the potential benefit they can gain (Ross et al. [2019]). Our conversations with Compassion Tanzania reveal that they select children in adverse conditions to lift them out of poverty.[11] Our STEM intervention was conducted in the Tabora region of

[11] The criteria for sponsorship considered by Compassion International in Tanzania are as follows: 1) children

Tanzania. In Tabora, Compassion has approximately 2700 sponsored children. These children are required to attend the Compassion centers every Saturday to maintain their sponsorship status.[12] These regular visits resemble weekly church activities, including reading the Bible, chanting hymns, and learning to dance or draw.[13]

Tabora is in the central-western part of the country, and its GDP per capita ranked 19th out of 23 regions in 2018 (National Bureau of Statistics [2019]). In Table 1, we compare the household characteristics of our study sample with representative samples from the Tabora region and Tanzania. It shows the shares of children who have a father in the households were lower for our study samples than the representative samples of Tabora or Tanzania. This is expected because one of Compassion's criteria is whether the children are from single-parent households. However, in terms of parental education and household wealth, our study sample is positively selected. Parents' educational level is higher, and household wealth tend to be higher for study samples. This could be because the students in our sample belong to the Tabora Urban district in the Tabora region, which is an urban area compared to the rest of districts in the Tabora region. Along the same line, sponsored children's educational aspirations tend to be high. Our data shows that students from Compassion in the Tabora region have high baseline educational aspirations with more than 80% of the respondents wanting to pursue higher education. This is partly due to the sponsorship program, which includes support for school materials and relationship building with sponsors (Ross et al. [2019]).[14]

To provide a new educational opportunity to sponsored children, Compassion partnered with e3empower, a local social enterprise specializing in quality science education and information and communication technologies (ICTs) for youth in developing countries. Compassion supported local centers in the Tabora region of Tanzania to provide this educational oppor-

live within 30 minutes from the center by foot, 2) children are from a disadvantaged background (e.g., single-parent household, orphaned), and 3) parents or guardians agree to receive sponsorship. Children leave the sponsorship program when their households' financial situation improves or when parents or guardians no longer want their children to participate in the program. The latter case often involves children staying at home to take care of younger siblings or when the child begins to work to support the family.

[12]While they are encouraged to aim for good grades in school, their sponsorship is not taken away if they fail.

[13]These activities are similar across all centers.

[14]For a more detailed description of the Compassion program, see Wydick et al. [2013] and Ross et al. [2019].

tunity devised by the e3empower staffs. Consequently, computer labs were installed at three centers, which became the main sites for the STEM boot camps offered to sponsored children from 12 local centers in the Tabora region. The locations of Compassion centers in Tabora for this study are shown in Figure 1.

3.2 Intervention

The STEM boot camp was a 5-day long extracurricular education program designed to provide hands-on experiences in STEM subjects. The boot camps took place at three Compassion centers with computer labs, from November 12 to 30, 2018 (Figure 2).[15][16] The program was conducted from 8:30am to 6:00pm for five days, right after the school instruction ended in the region. Table A2 shows the schedule of the boot camps. The camps consisted of an introduction to computers, sessions to learn coding skills, sessions to assemble robots, and sessions to create a simple app with program coding (Figure 3).[17] Many of these activities were group activities, usually with three students per group. Assignments were performed during the sessions; students used block coding to move robots along a line or set up a home alarm system using Arduino, an open-source hardware and software coding program. Their progress was reported during the class.[18] As the STEM boot camp contents require knowledge of basic arithmetic skills, this program was offered to children who were going to complete or had completed primary school (Standard 7 to Form 3 students), which corresponds to ages 13 to 16 years. The sessions were taught by technological experts from e3empower, which has provided STEM education programs in Tanzania since 2015. Students who had otherwise very little knowledge in computers or other ICT skills were trained to code and program using

[15]Three centers are AICT Kiete (Site 1), Moravian Chemchem (Site 2) and Anglican Isevya (Site 3).

[16]For secondary public schools in Tanzania, the general schedule of the academic year is from January to May, summer break in June, and another semester from July to the end of November or the first week of December. Form 2 and 4 students take national exams from the middle of November to the end of November. There are also two-week breaks in April and September, respectively. At Tabora, Form 1 and 3 students were invited during the second week of November to attend the boot camp, because they did not go to school, while Form 4 students were taking exams. Form 2 students were invited in consequent weeks, after taking national exams in the second week and after their break began.

[17]Compassion selected roughly an equal number of boys and girls to the boot camp. While the teaching staff did not address issues on gender equality during the STEM camp, Compassion generally has a more gender equality-friendly environment than a typical public school setting.

[18]All the assignments were in-session and there was no homework.

computers and build simple robots through the STEM boot camp program.

3.3 Baseline and endline data collection

Baseline surveys were conducted in early November 2018, just before STEM boot camps were offered (Figure 2). Students were unaware of the STEM bootcamp during the baseline surveys, and their willingness to participate was not asked. After the survey was conducted, students were informed about the program and whether they had been selected. The endline survey was run from mid-December 2018 to January 2019, two to ten weeks after completing the boot camps, depending on the participation week. We aimed to conduct a follow-up survey after one year, but could not do so due to COVID-19. For both boot camp participants and non-participants enrolled in the Tabora Compassion Centers, baseline and endline (paper-based) surveys were conducted at the centers they regularly visit on Saturdays. All surveys were conducted in Kiswahili, translated from the English version. The survey consisted of five broad categories: 1) basic information about the interview date and interviewee, 2) socioeconomic background information including parental education and household wealth-related measures, 3) students' aspiration including future education goals, preferred majors, and labor market-related outcomes, 4) self-esteem, and 5) gender attitudes. The questions were asked in a similar manner as in previous studies on development economics, especially the Demographic Health Survey conducted by the United States Agency for International Development (USAID).

Regarding gender attitudes which is the main outcome of interest, our survey consists of 11 questions regarding gender views on equal participation in domestic work, economic activities, and education. It also contains two questions on whether a respondent thinks men are superior to women in terms of making political leaders or business executives. These questions are comparable to the questions on gender norms available in the World Value Survey (Inglehart et al. [2014]) and Dhar et al. [2020]. Using answers from these questions, we construct a gender attitude index by applying a principal component analysis (PCA).[19]

[19] The index is based on the first component. The principal component loadings are shown in Table A3. The table shows that the principal component is mostly related to the attitudes on women's role in households, university education, and political participation.

However, alternative definitions of gender attitude index are possible. As robustness checks, we use two alternative definitions as well: (1) a simple average of the answers to 11 questions; and (2) an index following Anderson [2008]'s method.

4 Participant selection and empirical strategies

Compassion International in the Tabora region selected 180 participants among 829 enrolled students (Standard 7 to Form 3) for the STEM boot camp. The remaining students constituted the control group.[20] Using their best judgement, each center randomly selected students in a casual manner, after considering the class year and their attendance at the centers. To test the random assignment, we evaluate the balance between baseline characteristics of participants and non-participants, conditional on school year and religion, which is a proxy for attendance because Compassion is a Christian faith-based organization.

The summary statistics in Table A4 show baseline characteristics of participants and non-participants for the sample available in both baseline and endline surveys. As expected from the conversations with Compassion, Christian students were oversampled, and there were more Form 2 students among the participants. However, there were other imbalances as well. For example, STEM participants had lower average study hours for math and science subjects. This suggests that it is not the case that students who were more motivated in STEM education were selected. Another imbalance is the more progressive gender views of STEM participants, which is partly explained by religion and grades (in the last two columns in Table A4). We control for baseline gender attitudes to account for any differences in pre-treatment outcomes. [21]

With these imbalances in observables, we cannot entirely rule out the possible imbalances in the unobservables. Therefore, we do not interpret the STEM intervention as a completely

[20] The non-compliance rate was not zero, yet small at 1 percent, consisting of 9 out of 829 students. Among the students offered to participate in the program, 6 students did not participate, and 3 students – who were not offered – participated (less than 0.5 percent of 649 students not offered to participate). Most of the students gladly accepted the offer to participate in the program. The only challenge was scheduling conflicts since they had responsibilities at home or elsewhere.

[21] Although the average gender attitude index is higher for the treated group, the supports are well overlapped between the participant group and the non-participant group.

exogenous variation. Instead, we treat it as a *quasi-exogenous* variation and take a selection on observables (CIA) approach; we assume that STEM participation is exogenous conditional on rich baseline characteristics including pre-treatment gender attitudes. Although it is not possible to directly test CIA, we use methods from Oster [2019] and Altonji et al. [2005] to assess the degree of selection on unobservables that explains away the results.

Hirano and Imbens [2001] suggest that selection on observables approach is appropriate when much information about pre-treatment is available. Our extensive list of variables in the baseline includes detailed information on households, educational aspirations, personalities, and pre-treatment outcomes. Given these rich baseline characteristics combined with the information from the field that the scope of self-selection is limited, assuming exogeneity of the treatment conditional on baseline covariates is reasonable. To gauge the extent to which unobservables play roles in the given regressions, the method from Oster [2019] combines information from coefficient stability and movements in R-squared. We conduct this analysis to better understand the degree of selection on unobservables.

With the CIA, we use inverse probability weighting (IPW), exploiting the propensity scores (Imbens [2000]) to control for baseline characteristics. As exact matching is difficult given the high dimension of baseline characteristics, we weight the observations using the inverse of the propensity score to make the treatment and control groups comparable. Dehejia and Wahba [2002] show that estimators using propensity scores work well when there is substantial overlap in propensity scores between the treatment and control group. They also show that these estimators produce similar estimates to the benchmark experimental estimates when baseline characteristics, and importantly pre-treatment outcomes, are balanced after matching on propensity scores.[22] Figure A1 shows that the propensity scores for the two groups overlap well in most ranges, although the distribution of propensity score for the treated group has a higher propensity to be treated on average.[23] Moreover, Table 2 shows that after weighting,

[22]Dehejia and Wahba [2002] use propensity score matching. We use inverse probability weighting using propensity scores as the main specification but conduct propensity score matching as a robustness check (Table A10).

[23]The observations that do not have overlap are less than 5% of the samples. We repeat our analysis excluding samples in the non-overlap range as a robustness check and find very similar results.

the two groups are more comparable with the p-value of 0.99 for the tests of the equality of the two groups. Notably, baseline gender attitudes are similar between the treated group and control group after the weighting.

We use the following model to estimate the impact of the STEM program with the IPW:

$$Y_{ict} = \alpha + \beta STEM_{ic} + \delta Y_{ic,t-1} + \theta X_{ic,t-1} + \gamma_c + \epsilon_{ict} \tag{1}$$

where i, c, and t indicate individual students, Compassion centers, and time of the survey, respectively. We observe the outcomes, represented by the variable Y_{ict}, before and after the STEM education. The main outcome of interest is the index of gender attitudes, but we also investigate education-related and labor-market-related outcomes. $t-1$ indicates the outcome before the treatment, and t indicates the outcome after the treatment. The vector $X_{ic,t-1}$ comprises control variables, including all the variables in the baseline survey. γ_c is a set of center fixed effects. The propensity score is estimated using $X_{ic,t-1}$ and γ_c, and weights from the propensity scores are applied in each regression.

Finally, there was attrition in the sample between the baseline and post survey. 81% of students were surveyed in both waves, but we could not follow up with the rest. In particular, the non-participant group was less likely to take the follow-up survey. Table A5 shows differential attrition rates for the post-survey, with STEM participants having a lower attrition rate. We investigate the factors influencing this attrition pattern in the same table. Our investigation shows that students who attritted (mostly from the non-participant group) have more regressive gender views and are less motivated about educational goals. This suggests that our estimates would understate the true effects, providing conservative estimates. To examine the robustness of our results to the differential attrition, we use two methods. First, we report the upper and lower bound of the estimated effect using Lee [2009]'s sharp-bound estimators, which trim the highest or lowest values of the outcome variable (e.g., gender attitude index) from the STEM samples, the less-attritted study arm. This gives us a lower bound and an upper bound, respectively. Second, in the spirit of Horowitz and Manski [2000], we replace missing values with the 75 percentile and 25 percentiles of the outcome variable, providing the lower and upper bounds, respectively. Although we provide the lower bound, it is unlikely because attritted samples in the non-participant group have more regressive gender views.

5 Results

We present the results in this section. Section 5.1 shows the main results on gender attitudes and discusses potential concerns due to experimenter demand effects. In Section 5.2, we examine possible mechanisms that explain the main effects, and present a decomposition analysis. Section 5.3 discusses role model effects and peer effects on gender attitudes. Section 5.4 presents the results on behavior and long-term preferences and section 5.5 discusses the potential fade-out of the effects.

5.1 Gender attitudes

Table 3 presents our main results of the effect of STEM boot camp on gender attitudes, using a gender attitudes PCA index as the outcome variable. Conditional on baseline gender attitudes, column (2) adds controls for an extensive list of baseline characteristics, including pre-treatment self-esteem and personalities, labor market aspirations, and household characteristics such as parents' age and education, and household wealth indicators (e.g., access to electricity). Reassuringly, the coefficient in column (2) is hardly different from the one in column (1). Moreover, this coefficient stability is accompanied by a substantial increase in the R-squared, suggesting a limited scope of selection (Oster [2019]). The magnitude of the treatment effect and the pattern are also similar when we use an alternative definition of gender attitudes index, such as simple average index and index created following Anderson [2008] (Table A6).[24]

To understand the degree of selection on unobservables, we conduct an analysis outlined in Oster [2019]. This approach considers both coefficient stability and R-squared movements to address the possibility of selection on unobservables. Including uninformative controls, which

[24]The index used in Table A6 is created following the method in Anderson [2008] and Dhar et al. [2020]. To use this method, we create a binary variable where the most regressive view is 0 and the rest are 1. Then, we construct weights from the inverted covariance matrix of all the normalized variables in the gender survey and obtain the weighted sum as the index.

does not change R-squared much, would give stable coefficients even when a substantial bias exists. Therefore, coefficient stability has to be interpreted relative to R-squared movements. The Oster [2019] method allows us to obtain the relative degree of selection on observed and unobserved variables. Using this method, we find that after conditioning on center fixed effects and baseline gender outcomes (i.e., column (1) in Table 3 as the baseline model), the selection on unobservables would need to be 3.7 times as important as the observables to produce a treatment effect of zero.[25] In the literature, the reasonable cutoff for robustness is suggested to be 1, the same relative degree of selection of the observables and the unobservables (Altonji et al. [2005] and Oster [2019]).[26] Our number 3.7 is notably larger than this cutoff, assuring us that the results are not entirely driven by the selection on unobservables.[27]

Column (3) in Table 3 shows our preferred estimate using IPW. If the OLS controls the covariates in a flexible way, the result from IPW should not be very different from the OLS result (Angrist and Pischke [2008]). Our results show that this is the case. The coefficient in column (3) with the IPW model is similar to the coefficient in column (2) even though the size is slightly smaller. The result suggests that the gender attitudes of the STEM participants become more progressive by 0.16 standard deviations compared to the non-participants. Although our intervention does not directly aim to change gender attitudes but rather emphasizes hands-on experiences to learn STEM, the magnitude of the estimate is comparable to findings by Dhar et al. [2020], who conducted a school-based intervention program in India targeting gender attitudes.[28]

We implement two bounding exercises following Lee [2009] and Horowitz and Manski [2000] to address potential selection bias due to attrition. First, Lee's sharp bounds give us 0.083 (standard error (SE): 0.096) and 0.581 (SE: 0.098) as lower and upper bounds. Second, the

[25]The method from Oster [2019] is to evaluate the degree of selection on unobservables using information obtained from observables.

[26]One of the supporting arguments for this cutoff is that data collection of researchers is focused on the most important control variables.

[27]The results on the degree of selection for all outcomes are presented in Table A7.

[28]The interventions in Dhar et al. [2020] provided a gender equality education for approximately 20 hours over two and half years, and found that short-term gender attitudes changed by 0.179 standard deviations. Although our intervention was implemented in a shorter period of time, the total hours of intervention were over 45 hours, which was at least twice as much as the total time in Dhar et al. [2020].

upper and lower bounds calculated by replacing missing values with the 75 percentile and 25 percentile of the outcome variable are 0.157 (SE: 0.078) and 0.266 (SE: 0.079), respectively (Table A8). These numbers suggest that our results are robust to adjustments for attrition. Moreover, given that our attrition patterns show that attritted students in the non-STEM group have more regressive gender views, the lower bounds are unlikely. These numbers suggest that it is very unlikely that our results are driven by bias from the attrition.[29]

Next, we investigate heterogeneous effects by gender in Table 4.[30] We expect that our STEM intervention is likely to generate a differential effect on gender attitudes by gender, as discussed in Introduction. The results indeed show very different estimates for boys and girls; the STEM intervention increases the gender attitudes index of girls by a large amount (0.29 standard deviations), whereas the impact on boys is small at 0.035 and statistically insignificant.[31][32] Such finding indicates that positive effect on gender attitudes shown in Table 3 is mostly driven by the change in girls' attitudes. The absence of any effect among boys is consistent with previous research, reporting that boys often face difficulties in confronting prevailing gender norms due to social stigmatization (Kågesten et al. [2016]). The larger impacts on female students are also consistent with previous findings in China, showing an educational expansion policy influenced female gender attitudes (Du et al. [2020]). Our results suggest that exposure to new knowledge and a different domain of learning and work can be effective in changing gender attitudes, for girls who have limited opportunities relative to boys.[33] Given that we find a strong effect among girls only, we present and discuss our results focusing on girls. The results for boys are presented in the Appendix.

To further investigate which component of the gender attitudes the participants show im-

[29] There is a potential concern for spillover, if the participants returned to their school after the bootcamp and communicated with non-participants at school. As a result, the endline treatment effect might be contaminated. However, if there is any such effect from treated students' interacting with untreated students, the estimates would be biased downwards.

[30] The summary statistics by gender is presented in Table A9.

[31] We find similar results using propensity score matching (Table A10).

[32] We cannot reject the hypothesis that the treatment effect is the same across different camp sites or different camp weeks.

[33] Dhar et al. [2020], which provided interventions directly addressing gender equality, shows larger impacts on boys than girls, especially in the long run. For boys, who may take more opportunities for granted, interventions which directly aim to change gender attitudes appear to be more effective.

provements for girls, Figure 4 shows the estimates on disaggregated outcomes of the gender attitudes. Each outcome is converted so that a higher value represents a more progressive gender attitude. Therefore, positive coefficients indicate that the STEM boot camp made students have more progressive gender views. The STEM program shows a measurable positive impact on several dimensions of gender attitudes for girls. Among those, we find the largest impact on the traditional gender view that women should not earn more money than their husbands. Girls in the treated group disagreed more with this notion than the girls in the control group after the STEM intervention. The STEM boot camp exposed students to new skills, including computer coding and robot building, which may have affected the perceived earning potentials. This may ultimately have led to more progressive views on this specific outcome related to earnings for girls. We explore this channel more in-depth in the mechanism section. Girls in the treated group also showed fairly large improvements in gender attitudes on women's role, political participation, and tertiary education opportunities. These outcomes are related to educational opportunities and more active social participation, aligning with our intervention. However, we find little change in the traditional attitudes towards family structures (e.g., polygamy and wives' obedience to husbands).

5.1.1 Potential concerns due to experimenter demand effects

One potential issue regarding our estimate is the potential bias due to the experimenter demand effects (EDE). Under the EDE, subjects may change their behavior due to what they consider appropriate or socially desirable. Although our intervention does not directly address gender equality issues, STEM participants may have become more aware of socially appropriate responses even though they did not actually change their views or behaviors. Our baseline questionnaires about the big five personalities include questions that can be used to measure social desirability. We construct a social desirability index by applying the PCA using three baseline questions ("I feel little concern for others", "I insult people", and "I take time out for others") from the big five personalities, and include an interaction term between the index and the treatment dummy. These types of questions can be used to measure subjects' tendency to give appropriate answers and to test whether social desirability or EDE is the main driver for our results (Dhar et al. [2020]; Edmonds et al. [2020]).

Table 5 indicates that the higher the index (i.e., those who have a higher tendency to give socially appropriate answers), the higher the gender attitude indexes are. However, the interaction term between the treatment dummy and the dummy for the above-median social desirability index is small and statistically insignificant. Therefore, the treatment positively affects gender attitudes regardless of the underlying tendency of girls to give socially desirable answers.

5.2 Inspecting mechanisms

Our findings in Section 5.1 indicate that the intervention induces progressive gender attitudes among girls but not boys. This section explores a few channels to understand better the positive impacts of STEM boot camps on gender attitudes among girls.

Our intervention may affect gender norms through several channels. Specifically, the hands-on experiences and exposure to technologies for female students may boost their self-esteem, resulting in a change in gender attitudes. The intervention also exposes girls to STEM learning, which may increase their interest in STEM subjects, potentially affecting gendered stereotypes. Finally, the intervention may alter girls' perceived earnings in labor markets and career aspirations. In the following section, we first check whether outcomes related to these dimensions are affected by the treatment and then discuss the extent to which they explain the treatment effect on gender attitudes through a decomposition analysis.

5.2.1 Self-esteem

The first channel we investigate is self-esteem. Lower self-confidence in women is suggested to be one of the reasons that explains existing gender gaps in STEM choices (Niederle and Vesterlund [2010]). Like elsewhere, girls in Tanzania tend to perceive that STEM subjects are difficult (Semali and Mehta [2012]). Our STEM intervention with a strong emphasis on hands-on experiences and exposure to technologies may be effective at changing this perception, which may boost the girls' self-esteem. Increased self-esteem among girls may make gender attitudes more progressive, if they think they can do well in a wider range of fields.

Our survey contains three questions related to self-esteem, adopted from the Rosenberg self-esteem scale (Rosenberg [1965]): (1) overall satisfaction with oneself, (2) considering oneself having a number of good qualities, and (3) considering oneself doing things as well as most other people. Using these measures, we construct a single self-esteem index using PCA. The impacts of the STEM intervention on self-esteem are presented in panel A of Table 6. The intervention improves the self-esteem index among girls by a large degree, about 0.3 standard deviations. This finding is consistent with the previous studies showing that a robot education program for high schoolers in the U.S. and elementary school curriculum incorporating engineering aspects boosted girls' confidence in technology and interests in STEM (Nourbakhsh et al. [2005]; Rogers and Portsmore [2004]). This result suggests that self-esteem may be a channel that explains the positive impacts of STEM intervention on gender attitudes.

5.2.2 Interest in STEM subjects

The next possible channel is a change in interests in STEM subjects. Given our intervention provides an educational opportunity focused on STEM, the intervention may boost interests in learning STEM among girls. STEM is typically considered to be a male domain (Guimond and Roussel [2001]), and women are underrepresented in the STEM fields (Schwab et al. [2017]). Miller et al. [2015] use data from 66 countries and find a strong association between women's representation in science and the gender-science stereotypes. A strong association between gender-science stereotype and career choice (Makarova and Herzog [2015]) suggests that the boosted interest in STEM subjects could change the gender attitude index, which includes measures such as the perception on women's role as homemakers and perception on promoting equal university education opportunities for boys and girls.

Panel B in Table 6 reports the effects of STEM boot camps on the outcomes related to interest in STEM subjects. Using variables asking about favorite subjects, we create an indicator variable that takes value 1 if a student lists math or science as her favorite subject.

Similarly, we create binary variables indicating whether the student's second favorite subject and the third favorite subject are math or science.[34]

Panel B shows that the treatment increases the likelihood of choosing math and science subjects as the first favorite subject by 13.2 percentage points (23 percent of the mean of the control group), and increases the likelihood of picking the second favorite subject as math or science by 16.4 percentage points (37 percent of the control group mean). The treatment decreases the likelihood of choosing math or science as the the third favorite subject by 7.5 percentage points, smaller in magnitude than the results of the first two favorite subjects. This indicates that the increased level of interest in math and science subjects does not merely come from the students who escalated this interest from the third favorite to the second or the first favorite subject. Overall, the results in panel B suggest that the treatment mostly attracts students who previously had interests in subjects other than math and science to become interested in STEM subjects. Meanwhile, it also reinforces the interests among students weakly interested in STEM subjects before the intervention.

5.2.3 Expected performance and aspirations in labor markets

The results so far suggest that the treatment shapes girls' gender attitudes, accompanied by increased self-esteem while triggering interests in STEM subjects. Another potential channel is through an improvement in the perceived labor market opportunities. Jensen [2012] shows that expanded labor market opportunities increase career aspirations and change women's attitudes towards work and having children. Our intervention provides skill training on technologies, and these skills are likely to affect the potential earnings girls perceive. This may ultimately affect gender attitudes, especially the dimensions that relate to income-generating activities.

Panel C in Table 6 reports the results on students' expectations on relative earnings. In the survey, students listed their expected earnings compared to others on a 1 to 10 scale, where 1 indicates the bottom 10 percent and 10 indicates the top 10 percent of the income distribution.

[34] The survey asked the student to pick the top three favorite subjects in order, out of eight subjects. The list of the subjects in the questionnaire is: 1. Kiswahili, 2. Mathematics, 3. Science, 4. Geography, 5. Civics, 6. History, 7. English language, and 8. School sports.

The results show that the treated girls have a higher expectation of their earnings at age 25 by 0.92, roughly 14 percent of the control group girls on average. They exhibit a slightly modest view on their expected earning at age 35, roughly 7.4 percent higher than the control girls' average. Although the participants are still in school at the time of the survey and we do not have a direct measure of their long-term labor market outcomes, the results overall can be interpreted to have boosted students' confidence in their future performance in the labor market due to their experience at the STEM boot camp.

In panel D, we find that the treatment does not significantly influence career aspirations, although the sign is as expected.[35] Given the strong impacts of the STEM intervention on perceived earnings, labor market opportunities may be the main driver for the positive impacts on gender attitudes.

5.2.4 Decomposing treatment effects by channels

Considering the potential channels discussed above, we quantify their contribution to the main effects we find for girls using a decomposition method described by Heckman et al. [2013] and Gelbach [2016]. Specifically, we estimate the following model:

$$m^j_{ict} = \alpha^j + \beta^j STEM_{ic} + \delta^j Y_{ic,t-1} + \theta^j X_{ic,t-1} + \gamma^j_c + \epsilon^j_{ict} \tag{2}$$

where m_{ict} is a potential channel discussed in sections 5.2.1-5.2.3, and the other variables are as described in Equation 1. We can re-write Equation 1 in a long-form as follows:

$$Y_{ict} = \alpha^+ + \beta^+ STEM_{ic} + \delta^+ Y_{ic,t-1} + \theta^+ X_{ic,t-1} + \gamma^+_c + \sum_j \kappa^j m^j_{ict} + \epsilon^+_{ict} \tag{3}$$

where the contribution by channel j can be calculated as $\frac{\hat{\kappa}^j \hat{\beta}^j}{\hat{\beta}}$.

Figure 5 summarizes the results. We find that among the possible channels that we test, increased expected labor earnings explain the result by the largest degree, 16 percent. This

[35] We perform a PCA to create an index for career aspirations. We adopt the questionnaires from Gregor and O'Brien [2015]. The questions are included in Table A11.

result is also consistent with the finding in Figure 4, which shows that among the possible dimensions of gender attitudes, the treatment has the largest impact on updating the traditional belief that women should not earn more than their husbands.

Additionally, the boosted self-esteem explains the result by the second-largest extent, 9.4 percent. Increased interests in STEM subjects and career aspiration, while they are plausible channels, explain a relatively small extent of the treatment effect, 4.1 percent and 1.5 percent, respectively. Together, we find that the channels discussed earlier explain roughly 31 percent of the treatment effect.

5.3 Role models, peer effects, and gender attitudes

This section explores role model effects and peer effects through different sex-composition and peers' pre-program gender attitudes. We leverage variations in the teacher gender composition and peer characteristics based on the participating camp weeks and sites.[36]

Previous research suggests that a female role model can influence gender attitudes. Evidence from the U.S. universities shows that female role models affect major and class choices, especially in the STEM fields (Porter and Serra [2020]; Carrell et al. [2010]). Female teachers in middle school classrooms and a female role model in movies improve girls' academic performances (Lim and Meer [2017]; Riley et al. [2017]). Recent evidence shows that female role models positively influence the gender attitudes of boys and girls in Somalia (Kipchumba et al. [2021]). As such, the female role model or teacher's gender may have an independent effect on gender attitudes or it may affect the treatment heterogeneity by accentuating (or weakening) the treatment effect.

While our study was not designed to test the impacts of teachers' gender, there were variations in teacher gender in the STEM boot camps depending on the participating weeks and sites. Table A12 shows the composition of teachers. In Table 7, we examine whether the positive treatment effect is mainly driven by those girls exposed to a female instructor using the fact that a subset of girls had a female instructor. The interaction term between the

[36]As mentioned earlier, we do not find evidence that the treatment effects differ across different camp sites or different camp weeks.

treatment and having a female instructor at the STEM boot camps is positive but statistically insignificant. Thus, we do not have enough evidence that STEM boot camps impact gender attitudes through a female role model. Boys were also not differently affected by having a female teacher.

Next, we examine the influence of peers. Interactions with students of the opposite sex while solving STEM problems or exercises may have altered girls' gender attitudes (Markussen and Røed [2017]; Brenøe and Zölitz [2020]). Given the variations in the sex ratios of participants in different weeks and sites for the STEM boot camps, we can check whether the girls exposed to more male peers are more likely to change their gender attitudes.

Table A12 also shows the sex-composition during the STEM camp, which stayed constant during the 5-day camp period; there exist significant variations in male to female ratios ranging from 0.44 to 1.64. Column (2) in Table 7 presents the results by adding an interaction term between the treatment dummy and a binary variable indicating whether the student was exposed to a high female-to-male sex ratio setting during the STEM boot camps. The coefficient of the interaction term is close to zero, suggesting that sex-composition is not the main driver of our results. Further, we do not find a differential impact of the STEM intervention by gender compositions for boys.

We also explore another dimension of peer effects: whether the average peer's pre-program gender attitudes influences the STEM program's impact. To test this, we add an interaction between the treatment and a dummy for the above-median average peer's gender attitudes in the baseline. If these girls change their gender attitudes mainly by interacting with peers with progressive gender attitudes rather than by participating in the STEM program itself, the estimated coefficient on the interaction term would be large and positive, while the coefficient on STEM would be small and insignificant. The coefficient of the dummy in column (3) of Table 7 is positive, suggesting that a higher average of peers' gender attitudes positively affects the girls' gender attitudes. However, this coefficient is imprecisely estimated, and the coefficient on STEM is still sizable.[37] Interestingly, boys show improvements in gender attitudes if they

[37] We obtain similar results when we construct the interaction terms in columns (1) and (2) using a continuous

are surrounded by peers who tend to have higher baseline gender attitudes (Table A15).

5.4 Results on behavior and long-term preferences

This section examines whether the STEM boot camp leads to changes in behavior and in long-term preferences related to education, occupations, and family planning. Considering that the STEM boot camp is an educational opportunity, girls' sense of agency may be improved, and they may have a better outlook for their future (Edmonds et al. [2020]; Bandiera et al. [2020]). We take advantage of our survey questions on weekly study hours to evaluate girls' behavioral changes. For longer-term choices, although we do not observe actual choices of students, we have variables on their preferred choices for the future, including majors, occupations, and family structures. We investigate these variables to understand how their long-term preferences changed after the intervention.

Our findings in Table 8 show that the STEM boot camp influences not only girls' gender attitudes but also girls' behaviors and preferences on future choices related to education, marriage, and jobs. First, Panel A shows the impact of the boot camp on the amount of time spent by girls in studying. The girls in the treatment group spent significantly more time studying, 5 hours per week on average. Although they increased the largest amount of time (2 hours) studying STEM subjects, we also find spillover effects towards other subjects; girls also spent more time on other subject groups, including Kiswahili/English and geography/civics/history.

Then, Panel B shows whether girls in the treatment group aim to complete more years of education after the intervention. The treatment does not show a meaningful effect on improving educational goals, likely because most students in our sample (86 percent of the control group) already aspire to pursue higher education at baseline. However, among girls who intend to go to university, we find a significant impact on the preferred major in Panel C. Girls in the treatment group are more likely to opt to major in engineering or technology than girls in the control group by 18.9 percentage points, and they are less likely to choose other majors, business in particular. Given that improving technical skills is likely to yield

measure of sex ratio and average peers' pre-gender attitudes, instead of a binary indicator.

higher returns in the labor market, the treatment may also affect preferences related to the labor market, such as students' expected occupation. In return, the treatment may also affect marriage and parenthood preferences, given the strong relationship between the labor force participation and the opportunity cost of having a child.

Panel D shows that girls in the treated group exhibit a significantly lower likelihood of having an expected occupation such as clerical, sales, and services than non-treated girls. They also have a lower propensity to pursue an occupation that requires relatively low skills, such as agriculture, domestic service, and other manual jobs. In contrast, the treatment leads to an increase in STEM-related occupations (engineering, science, and technology sectors) and non-government professional/managerial jobs, although the effect is not precisely estimated. Lastly, Panel E shows the treatment effects on family-related outcomes. We find that although girls exposed to the STEM program desire to have fewer children by 0.26, roughly 7 percent of the mean, they do not necessarily want to delay their marriage. The result for the desired number of children is consistent with the finding that they expect to earn more and, as a result, would face a higher opportunity cost of having children.[38]

To summarize, although the treatment does not affect the aspiration for higher education, which was already quite high, it has a measurable effect on STEM-related educational outcomes. The program increased current and aspirational interests in STEM subjects, as seen by an interest in choosing STEM subjects as their preferred major at university. More importantly, this triggered interest in STEM subjects is consistent with students' behavioral change, as shown by the increased study hours in STEM subjects.

5.5 Potential for fade-out effects

Given that our post-survey was conducted 2-10 weeks after the intervention, one may question whether the positive effect on gender attitudes among girls is short-lived. We leverage

[38] This finding is consistent with the finding in Jensen [2012]. He shows that women in the villages with new employment opportunities in India not only expressed preferences on fewer number of children but also were actually less likely to have children during the study period.

the differences in the timing of the intervention and the post-survey to test this possibility.[39] We construct a binary variable indicating whether the students were surveyed before or after 33 days after the intervention, where 33 is the median of the difference in days between the intervention and the post-survey. To test whether the proximity of the survey to the intervention date affects the impacts of the intervention, we add an interaction term between the intervention and the above-median indicator of the days between the intervention and the post-survey. Table 9 presents the results. Although the impact size is slightly smaller for the students whose gap between the intervention date and the post survey is above the median, this difference is not statistically significant. Moreover, the positive effect remains significant for these students surveyed more than 33 days after the intervention. This shows that our findings are not temporary effects that exist only right after the intervention.

We look into insights from the literature for the persistence of the effects in the longer term. Dhar et al. [2020] investigate the impacts of a gender-equality education program in India on gender attitudes and found an impact of a size similar to our study. In their study, two endline surveys were conducted; the first endline survey took place 1-5 months after the intervention, and the second endline survey was collected two years after the intervention. They found very similar results in the first and the second endline surveys. The magnitude of the effect decreased by 10 percent over the two years, and strong positive impacts on gender attitudes remained. Kipchumba et al. [2021] provide suggestive evidence that the effects of female role models on gender attitudes persist in the Somali elementary school context. They evaluate the impact of a role model intervention where college students visit elementary schools as role models. They find a sizable impact (0.2 standard deviations) two years after the intervention although the impact was imprecisely estimated due to small sample size used to estimate the long-run effect. Although we cannot directly test our longer-term effects, these studies suggest that the impacts on gender attitudes may persist.

[39] This difference was arguably random; weather, road conditions and the availability of the enumerators affected the timing between the intervention and the post survey for the students. The difference is distributed from 12 to 75 days.

6 Conclusion

This study explores whether hands-on learning experiences focusing on STEM subjects reshape gender attitudes in Tanzania. Our results provide several implications. First, while there are many ways to improve gender attitudes, including the program that aims to directly reshape the gender attitudes (Dhar et al. [2020]), we find that the indirect approach of offering STEM education can be as effective in shaping gender attitudes among girls. Second, STEM education yields interests in STEM-related education, occupation, and higher aspirations in labor markets in the future among girls. The improved prospects in the labor market are also reflected in family decisions: reducing the number of desired children, suggesting a higher perceived opportunity cost of having children among girls in the treatment group. Our results suggest that a similar program that triggers interest in STEM subjects among adolescents may improve gender attitudes while providing them with adequate skills to partake in the digital era. While there has been a strong emphasis on STEM education in developed countries, it is still lagging due to a lack of quality teachers and learning resources in developing countries. For example, despite recent economic growth in sub-Saharan African countries, the share of high-skilled employment is 6% in the region, while the global average is 24% (Leopold et al. [2017]).

Lastly, one challenge that emerges from our study is that although boys benefit from the STEM program by having increased self-esteem and educational aspiration, the program does not improve boys' gender attitudes, who initially expressed less progressive gender attitudes than girls. Perhaps for boys, a more direct intervention targeting gender attitudes may be effective (Dhar et al. [2020]). Carefully designing STEM programs or other educational approaches that can improve gender attitudes for both boys and girls will be an important avenue for future research.

References

Abrevaya, Jason (2009) "Are There Missing Girls in the United States? Evidence from Birth Data," *American Economic Journal: Applied Economics*, 1 (2), 1–34.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn (2013) "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 128 (2), 469–530.

Almond, Douglas and Lena Edlund (2008) "Son-biased Sex Ratios in the 2000 United States Census," *Proceedings of the National Academy of Sciences*, 105 (15), 5681–5682.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber (2005) "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113 (1), 151–184.

Anderson, Michael L (2008) "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103 (484), 1481–1495.

Angrist, Joshua D and Jörn-Steffen Pischke (2008) *Mostly harmless econometrics: An empiricist's companion*: Princeton university press.

Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn (2020) "Negotiating a Better Future: How Interpersonal Skills Facilitate Intergenerational Investment," *The Quarterly Journal of Economics*, 135 (2), 1095–1151.

Auletto, Amy, Taeyeon Kim, and Rachel Marias (2017) "Educational Attainment and Egalitarian Attitudes toward Women in the MENA Region: Insights from the Arab Barometer," *Current Issues in Comparative Education*, 20 (1), 45–67.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman (2020) "Women's Empowerment in Action: Evidence from a

Randomized Control Trial in Africa," *American Economic Journal: Applied Economics*, 12 (1), 210–59.

Bauernschuster, Stefan, Oliver Falck, and Ludger Woessmann (2014) "Surfing Alone? The Internet and Social Capital: Evidence from an Unforeseeable Technological Mistake," *Journal of Public Economics*, 117 (C), 73–89.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova (2009) "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, 124 (4), 1497–1540.

Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan (2020) "Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women," *The Review of Economic Studies*.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015) "Gender Identity and Relative Income within Households," *The Quarterly Journal of Economics*, 130 (2), 571–614.

Brenøe, Anne Ardila and Ulf Zölitz (2020) "Exposure to More Female Peers Widens the Gender Gap in STEM Participation," *Journal of Labor Economics*, 38 (4), 1009–1054.

Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, and Iman Sen (2017) "Power vs Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial," Unpublished Manuscript.

Bulman, George and Robert W. Fairlie (2016) "Technology and Education: Computers, Software, and the Internet," *In Handbook of the Economics of Education*, 5, 239–280.

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

Compassion (2020) "2019 Annual Report," `https://annualreport.compassion.com/`.

——— (2021) “Compassion Tanzania,” `https://www.compassionuk.org/tanzania/`.

Crone, Eveline A and Ronald E Dahl (2012) “Understanding Adolescence as a Period of Social–affective Engagement and Goal Flexibility,” *Nature Reviews Neuroscience*, 13 (9), 636–650.

Dehejia, Rajeev H and Sadek Wahba (2002) “Propensity score-matching methods for nonexperimental causal studies,” *Review of Economics and statistics*, 84 (1), 151–161.

Del Carpio, Lucía and Maria Guadalupe (2018) “More Women in Tech? Evidence from a Field Experiment Addressing Social Identity,” IZA Discussion Papers 11876, Institute of Labor Economics (IZA).

Dhar, Diva, Tarun Jain, and Seema Jayachandran (2019) “Intergenerational Transmission of Gender Attitudes: Evidence from India,” *The Journal of Development Studies*, 55 (12), 2572–2592.

——— (2020) “Reshaping Adolescents’ Gender Attitudes: Evidence from a School-Based Experiment in India,” NBER Working Papers 25331, National Bureau of Economic Research, Inc.

Du, Huichao, Yun Xiao, and Liqiu Zhao (2020) “Education and Gender Role Attitudes,” *Journal of Population Economics*, 1–39.

Duflo, Esther (2012) “Women Empowerment and Economic Development,” *Journal of Economic Literature*, 50 (4), 1051–79.

Dynarski, Mark, Roberto Agodini, Sheila Heaviside, Timothy Novak, Nancy Carey, Larissa Campuzano, Barbara Means, Robert Murphy, William Penuel, Hal Javitz et al. (2007) “Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort. Report to Congress.,” *National Center for Education Evaluation and Regional Assistance*.

Edmonds, Eric V., Benjamin Feigenberg, and Jessica Leight (2020) "Advancing the Agency of Adolescent Girls," NBER Working Papers 27513, National Bureau of Economic Research, Inc.

Fairlie, Robert W. and Jonathan Robinson (2013) "Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren," *American Economic Journal: Applied Economics*, 5 (3), 211–40.

Gelbach, Jonah B (2016) "When do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics*, 34 (2), 509–543.

Gregor, Margo A and Karen M O'Brien (2015) "The Changing Face of Psychology: Leadership Aspirations of Female Doctoral Students," *The Counseling Psychologist*, 43 (8), 1090–1113.

Guimond, Serge and Lydie Roussel (2001) "Bragging about One's SchoolGgrades: Gender Stereotyping and Students' Perception of their Abilities in Science, Mathematics, and Language," *Social psychology of education*, 4 (3), 275–293.

Heckman, James, Rodrigo Pinto, and Peter Savelyev (2013) "Understanding the Mechanisms Through which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 103 (6), 2052–86.

Hirano, Keisuke and Guido W Imbens (2001) "Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization," *Health Services and Outcomes research methodology*, 2 (3), 259–278.

Horowitz, Joel L and Charles F Manski (2000) "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," *Journal of the American statistical Association*, 95 (449), 77–84.

Imbens, Guido W (2000) "The role of the propensity score in estimating dose-response functions," *Biometrika*, 87 (3), 706–710.

Inglehart, Ronald, Christian Haerpfer, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, Jaime Diez-Medrano, Marta Lagos, Pippa Norris, Eduard Ponarin, Bi Puranen et al. (2014) "World Values Survey: Round Six-Country-pooled Datafile Version," *Madrid: JD Systems Institute*, 12.

International Labour Office (2018) "Global Wage Report 2018/19: What lies Behind Gender Pay Gaps."

Jayachandran, Seema (2015) "The Roots of Gender Inequality in Developing Countries," *Annual Review of Economics*, 7 (1), 63–88.

——— (2020) "Social Norms as a Barrier to Women's Employment in Developing Countries," Working Paper, National Bureau of Economic Research.

Jensen, Robert (2012) "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *The Quarterly Journal of Economics*, 127 (2), 753–792.

Jensen, Robert and Emily Oster (2009) "The Power of TV: Cable Television and Women's Status in India," *The Quarterly Journal of Economics*, 124 (3), 1057–1094.

Kågesten, Anna, Susannah Gibbs, Robert Wm Blum, Caroline Moreau, Venkatraman Chandra-Mouli, Ann Herbert, and Avni Amin (2016) "Understanding Factors that Shape Gender Attitudes in Early Adolescence Globally: A Mixed-methods Systematic Review," *PloS one*, 11 (6), e0157805.

Kahn, Shulamit and Donna Ginther (2017) "Women and STEM," Working Paper, National Bureau of Economic Research.

Kane, Emily W (1995) "Education and Beliefs about Gender Inequality," *Social Problems*, 42 (1), 74–90.

Kipchumba, Elijah, Catherine Porter, Danila Serra, Munshi Sulaiman et al. (2021) "Infuencing Youths' Aspirations and Gender Attitudes through Role Models: Evidence from Somali schools,"Technical report.

Koch, Alexander, Julia Nafziger, and Helena Skyt Nielsen (2015) "Behavioral Economics of Education," *Journal of Economic Behavior & Organization*, 115, 3–17.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea (2012) "Soap operas and fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4 (4), 1–31.

Leaper, Campbell, Timea Farkas, and Christia Spears Brown (2012) "Adolescent Girls' Experiences and Gender-related Beliefs in Relation to Their Motivation in Math/Science and English," *Journal of Youth and Adolescence*, 41 (3), 268–282.

Lee, David S (2009) "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 76 (3), 1071–1102.

Leopold, Till Alexander, Vesselina Ratcheva, Saa Zahidi, and R Samans (2017) "The Future of Jobs and Skills in Africa: Preparing the Region for the Fourth Industrial Revolution," Tech. Rep., World Economic Forum, Geneva, Switzerland.

Lim, Jaegeum and Jonathan Meer (2017) "The Impact of Teacher–student Gender Matches Random Assignment Evidence from South Korea," *Journal of Human Resources*, 52 (4), 979–997.

Makarova, Elena, Belinda Aeschlimann, and Walter Herzog (2019) "The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations," in *Frontiers in Education*, 4, 60, Frontiers.

Makarova, Elena and Walter Herzog (2015) "Trapped in the Gender Stereotype? The Image of Science among Secondary School Students and Teachers," *Equality, Diversity and Inclusion: An International Journal*.

Malamud, Ofer and Cristian Pop-Eleches (2011) "Home Computer Use and the Development of Human Capital," *The Quarterly Journal of Economics*, 126 (2), 987–1027.

Markussen, Simen and Knut Røed (2017) "The Gender Gap in Entrepreneurship–The Role of Peer Effects," *Journal of Economic Behavior & Organization*, 134, 356–373.

McKelway, Madeline (2020) "Women's Employment in India: Intra-household and Intra-personal Constraints," Working Paper.

Miller, David I, Alice H Eagly, and Marcia C Linn (2015) "Women's Representation in Science predicts National Gender-Science Stereotypes: Evidence from 66 Nations.," *Journal of Educational Psychology*, 107 (3), 631.

Moen, Phyllis, Mary Ann Erickson, and Donna Dempster-McClain (1997) "Their Mother's Daughters? The Intergenerational Transmission of Gender Attitudes in a World of Changing Roles," *Journal of Marriage and Family*, 59 (2), 281–293.

MOEST (2019) "United Republic of Tanzania Education Sector Performance Report (2018/2019)."

National Bureau of Statistics (2019) "National Accounts Statistics, UNITED REPUBLIC OF TANZANIA," `https://www.nbs.go.tz/nbs/takwimu/na/National_Accounts_Statistics_Popular_Version_2018.pdf`.

National Center for Education Statistics (2018) "Annual Reports," `https://nces.ed.gov/programs/digest/d18/tables/dt18_218.10.asp?current=yes`.

Niederle, Muriel and Lise Vesterlund (2010) "Explaining the Gender Gap in Math Test Scores: The Role of Competition," *Journal of Economic Perspectives*, 24 (2), 129–44.

Nourbakhsh, Illah R, Kevin Crowley, Ajinkya Bhave, Emily Hamner, Thomas Hsiu, Andres Perez-Bergquist, Steve Richards, and Katie Wilkinson (2005) "The Robotic Autonomy Mobile Robotics Course: Robot Design, Curriculum Design and Educational Assessment," *Autonomous Robots*, 18 (1), 103–127.

Nuffic (2015) "Education System Tanzania Described and Compared with the Dutch System."

Oster, Emily (2019) "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 37 (2), 187–204.

Porter, Catherine and Danila Serra (2020) "Gender Differences in the Choice of Major: The Importance of Female Role Models," *American Economic Journal: Applied Economics*, 12 (3), 226–54.

Riley, Emma et al. (2017) "Role Models in Movies: The Impact of Queen of Katwe on Students' Educational Attainment," Working Paper, Centre for the Study of African Economies, University of Oxford.

Rogers, Chris and Merredith Portsmore (2004) "Bringing Engineering to Elementary School," *Journal of STEM Education: Innovations and Research*, 5 (3).

Rosenberg, Morris (1965) "Rosenberg Self-esteem Scale (RSE)," *Acceptance and Commitment Therapy. Measures package*, 61 (52), 18.

Ross, Phillip H., Paul Glewwe, Daniel Prudencio, and Bruce Wydick (2019) "Developing Educational and Vocational Aspirations through International Child Sponsorship: Evidence from Kenya, Indonesia, and Mexico," Miscellaneous Publications 302342, University of Minnesota, Department of Applied Economics.

Schwab, Klaus, Richard Samans, Saadia Zahidi, Till Alexander Leopold, Vesselina Ratcheva, Ricardo Hausmann, and Laura D Tyson (2017) "The Global Gender Gap Report 2017," World Economic Forum.

Semali, Ladislaus M and Khanjan Mehta (2012) "Science Education in Tanzania: Challenges and Policy Responses," *International Journal of Educational Research*, 53, 225–239.

Tanzania National Bureau of Statistics (2017) "Women and Men in Tanzania: Facts and Figures."

——— (2018) "Women and Men in Tanzania: Facts and Figures."

The United Republic of Tanzania President's Office (2016) "Pre-Primary, Primary and Secondary Education Statistics in Brief."

UN Statistics Division (2018) "Minimum Set of Gender Indicators," https://genderstats.un.org/#/countries.

UNDP (2019) "Human Development Reports: Gender Inequality Index (GII)," http://hdr.undp.org/en/content/gender-inequality-index-gii.

UNESCO (2009) "Needs Assessment Study of Tanzania's Science Education."

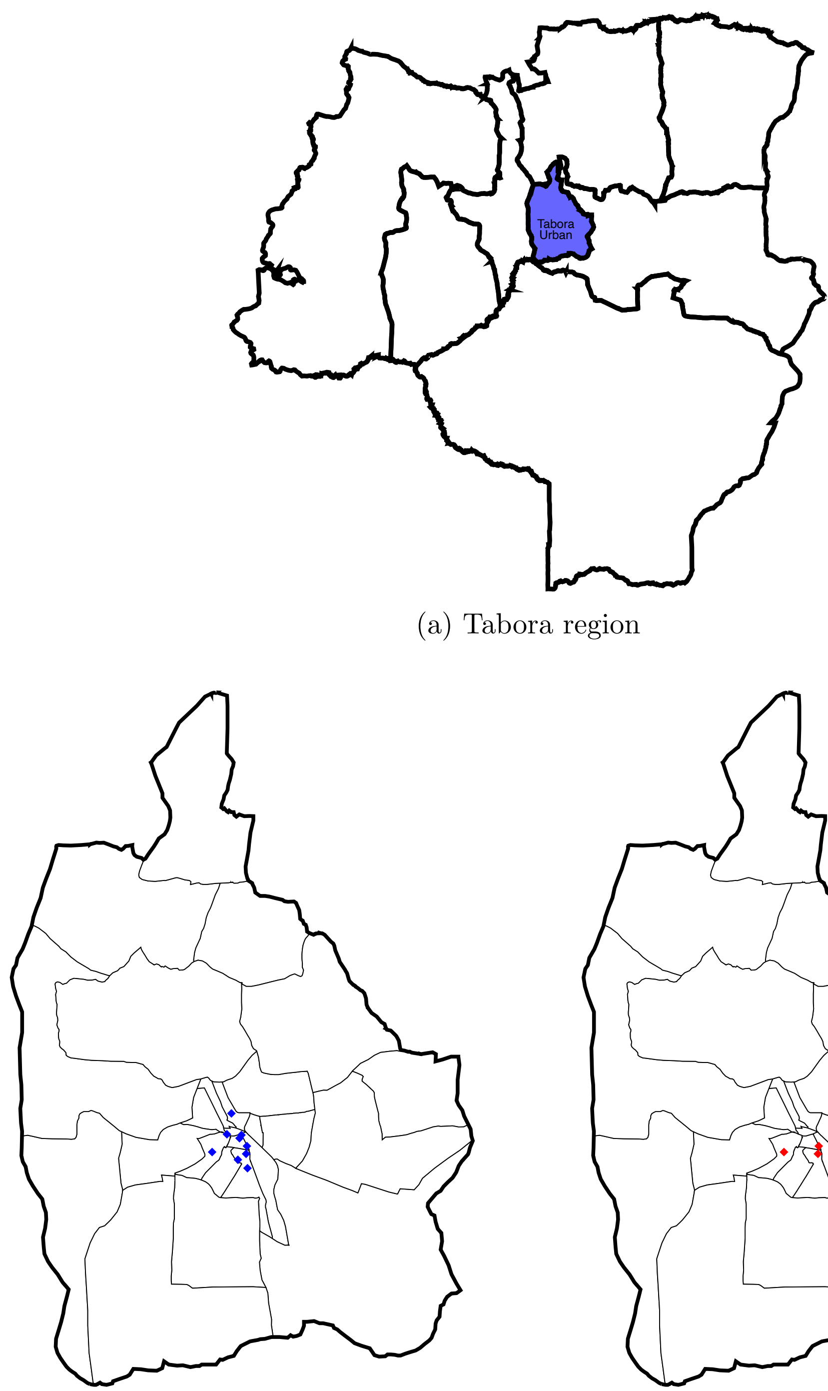
UNICEF (2017) "Education Fact Sheet."

World Bank (2018) "World Development Report 2018: Learning to Realize Education's Promise."

——— (2021) "World Bank Open Data," https://data.worldbank.org/.

Wydick, Bruce, Paul Glewwe, and Laine Rutledge (2013) "Does International Child Sponsorship Work? A Six-country Study of Impacts on Adult Life Outcomes," *Journal of Political Economy*, 121 (2), 393–436.

Figure 1: Compassion centers and intervention locations



(a) Tabora region

(b) Compassion centers in the Tabora Urban district

(c) STEM sites in the Tabora Urban district

Notes: The thick lines in (a) indicate the districts of the Tabora region. The shaded blue area indicates the Tabora Urban district. The thick lines in (b) and (c) indicate the Tabora Urban district and the fine lines in (b) and (c) indicate the wards in the Tabora Urban district.

Figure 2: Timeline

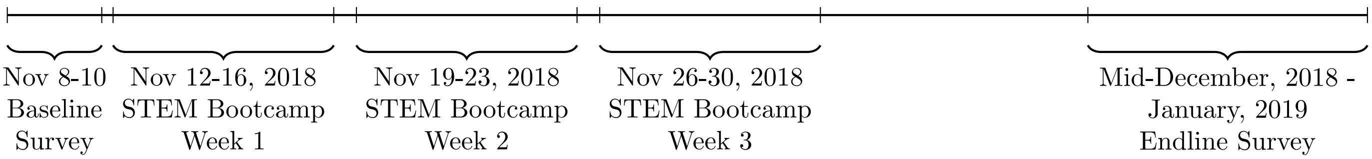
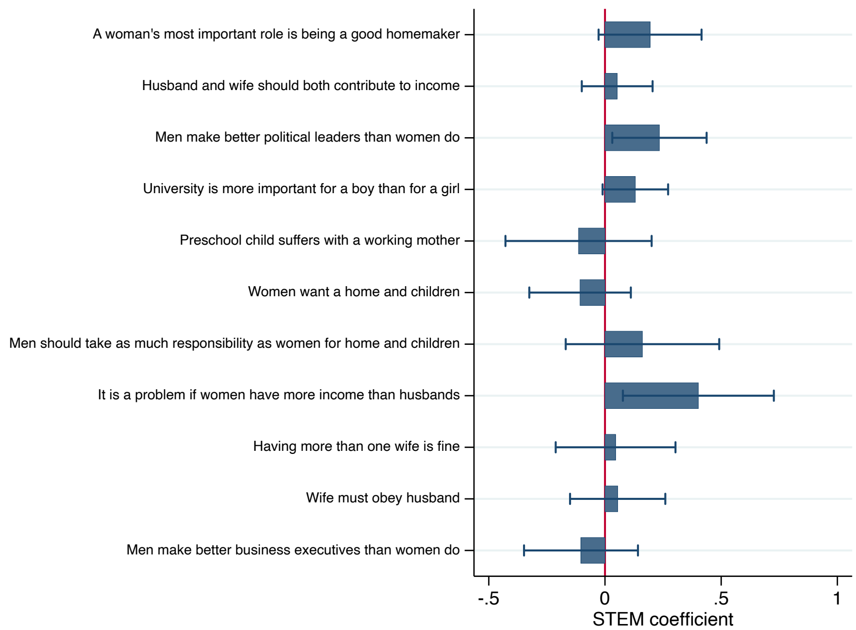


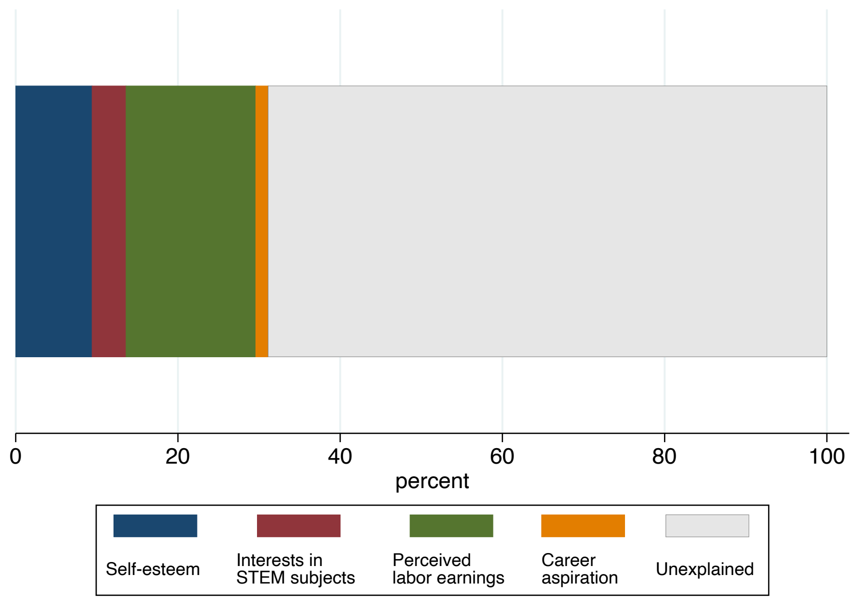
Figure 3: STEM interventions

Figure 4: The effects of STEM intervention on gender attitudes for girls



Notes: This figure presents the treatment effects on disaggregated gender attitude outcomes, and 95% confidence intervals. The answers are transformed so that higher number indicates more progressive gender views for each item. Therefore, the positive coefficients mean that gender attitudes became more progressive after the intervention for the corresponding item.

Figure 5: Decomposition analysis for possible channels



Notes: This figure presents the estimated decomposition of STEM intervention effects on gender attitudes. The effects on gender attitudes can be explained by self-esteem (9.4%), interests in STEM subjects (4.1%), perceived labor market earnings (16.0%), career aspirations (1.5%), and other factors.

Table 1: Summary statistics on household characteristics for sampled students

	(1) Sample mean	(2) Tabora mean	(3) Tanzania mean
Father in household	0.587	0.686	0.699
Mother in household	0.816	0.705	0.803
Father education: No schooling	0.033	0.096	0.102
Father education: Primary	0.371	0.720	0.747
Father education: More than primary	0.596	0.185	0.151
Mother education: No schooling	0.057	0.118	0.142
Mother education: Primary	0.463	0.861	0.763
Mother education: More than primary	0.481	0.021	0.095
Electricity or solar energy source	0.684	0.755	0.647
Number of rooms in households	3.169	2.514	2.983
Piped water source	0.724	0.090	0.498
Flush toilet	0.426	0.138	0.170

Notes: Column (1) presents the mean values averaged over all students in the study sample. Columns (2) and (3) are calculated using student samples who are in Standard 7 - Form 3 from Tanzania Household Budget Survey (HBS) 2017-2018.

Table 2: Summary statistics (with inverse probability weighting)

	STEM		Non-STEM		Diff		Diff (w/ center FE)	
	Mean	SD	Mean	SD	Coef.	SE	Coef.	SE
Age	15.032	1.609	14.930	1.530	0.102	0.174	0.120	0.164
Female	0.517	0.501	0.494	0.500	0.023	0.055	0.019	0.053
Grade:								
Standard7	0.209	0.408	0.256	0.437	-0.047	0.040	-0.043	0.037
Form 1	0.339	0.475	0.350	0.478	-0.011	0.051	-0.014	0.050
Form 2	0.256	0.438	0.240	0.428	0.016	0.044	0.018	0.045
Form 3	0.196	0.398	0.153	0.361	0.042	0.057	0.038	0.050
Religion:								
Christianity	0.695	0.460	0.641	0.476	0.054	0.051	0.048	0.048
Islam	0.305	0.460	0.359	0.476	-0.054	0.051	-0.048	0.048
Father in household	0.591	0.453	0.594	0.449	-0.004	0.050	-0.007	0.048
Mother in household	0.830	0.363	0.816	0.369	0.014	0.036	0.014	0.035
Father years of education	9.828	3.322	9.658	3.240	0.170	0.369	0.117	0.312
Mother years of education	8.841	3.024	8.702	3.395	0.138	0.334	0.101	0.293
Piped water source	0.735	0.437	0.722	0.445	0.013	0.048	0.011	0.044
Electricity or solar energy source	0.695	0.456	0.680	0.463	0.016	0.048	0.011	0.044
Flush toilet	0.457	0.494	0.425	0.490	0.032	0.056	0.028	0.055
Number of rooms in HHs	3.147	1.520	3.134	1.465	0.014	0.134	0.015	0.131
Computer use:								
Ever used a computer at school	0.216	0.398	0.172	0.359	0.044	0.056	0.038	0.050
Hours on math and science per week	6.325	6.251	6.027	6.579	0.298	0.795	0.301	0.712
Educational goal:								
Senior secondary or below	0.041	0.193	0.058	0.227	-0.018	0.019	-0.018	0.019
Technical/ vocational	0.153	0.357	0.111	0.306	0.042	0.054	0.038	0.047
Higher education	0.806	0.390	0.830	0.365	-0.025	0.055	-0.020	0.047
Expected occupation:								
Clerical, sales and services	0.013	0.109	0.017	0.124	-0.004	0.009	-0.005	0.009
Education sector	0.128	0.331	0.137	0.332	-0.009	0.033	-0.007	0.032
Engineering/science/technology sector	0.296	0.451	0.305	0.446	-0.009	0.046	-0.011	0.047
Health and social/community work	0.088	0.281	0.090	0.277	-0.002	0.026	-0.001	0.026
Professional/ managerial (government)	0.331	0.465	0.337	0.457	-0.005	0.050	-0.001	0.046
Professional/ managerial (non-government)	0.061	0.238	0.060	0.230	0.001	0.028	-0.001	0.026
Agriculture/domestic service/manual job	0.082	0.273	0.054	0.219	0.028	0.051	0.025	0.043
Desired marriage age (male)	27.439	5.006	27.559	4.526	-0.121	0.704	-0.186	0.704
Desired marriage age (female)	28.385	3.375	27.940	4.817	0.446	0.509	0.263	0.499
Standardized gender attitudes	0.129	1.156	0.112	0.992	0.017	0.182	0.020	0.166
Standardized self-esteem index	0.130	0.922	0.054	0.970	0.076	0.102	0.071	0.097
Big 5 personality:								
Extroversion	0.053	0.916	0.023	0.970	0.030	0.090	0.020	0.089
Agreeableness	0.097	0.882	0.043	0.952	0.055	0.089	0.040	0.089
Conscientiousness	0.091	0.859	0.044	0.951	0.047	0.085	0.041	0.085
Neuroticism	0.081	0.831	0.036	0.954	0.045	0.083	0.041	0.083
Openness	0.096	0.857	0.035	0.956	0.061	0.082	0.053	0.078
P-value for joint test of orthogonality					0.997		0.997	
Number of observations	174		494					

Notes: Significance levels: * 10%, ** 5%, *** 1%. The results in the last four columns are obtained from the regressions using an inverse probability weighting. The propensity scores for the weights are estimated using all available variables in the baseline survey. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Self-esteem index is constructed using a PCA using three questions from the Rosenberg's self-esteem scale.

Table 3: The effects of the STEM bootcamp on gender attitudes

	Dependent variable: Gender attitudes		
	(1)	(2)	(3)
	OLS	OLS	Inverse probability weighting
STEM	0.189**	0.183**	0.163**
	(0.079)	(0.084)	(0.077)
Control baseline gender attitudes	Yes	Yes	IPW
Control covariates	No	Yes	IPW
Control group mean	0	0	0
R-squared	0.128	0.240	
Observations	668	668	668

Notes: Significance levels: * 10%, ** 5%, *** 1%. Covariate variables include all variables available in the baseline survey and squared and cubic terms for continuous variables. Center fixed effects are included in all specifications. Propensity score for inverse probability weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Unweighted control mean and standard deviation are normalized to 0 and 1 for the PCA index.

Table 4: The effects of the STEM bootcamp by gender

	Dependent variable: Gender attitudes	
	(1)	(2)
	Males	Females
STEM	0.035	0.294***
	(0.119)	(0.101)
Randomization inference p-values	.815	.049
Control group mean	-.308	.318
Observations	339	329

Notes: Significance levels: * 10%, ** 5%, *** 1%. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

Table 5: Heterogeneous effects by social desirability index for girls

	Dependent variable: Gender attitudes
	(1)
STEM	0.279*
	(0.168)
STEM × Above median social desirability index	-0.068
	(0.249)
Above median social desirability index	0.298
	(0.182)
Control group mean	.318
Observations	329

Notes: Significance levels: * 10%, ** 5%, *** 1%. Social desirability index is constructed by applying a PCA using 3 baseline questions ("I feel little concern for others", "I insult people" and "I take time out for others") in big 5 personalities. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

Table 6: Potential channels for girls

Dependent variable	STEM	SE	Control mean
A. Self-esteem			
Standardized self-esteem index	0.296*	0.164	-0.060
B. Interests in STEM subjects			
1st favorite subject-math and science	0.132**	0.059	0.563
2nd favorite subject-math and science	0.164**	0.065	0.447
3rd favorite subject-math and science	-0.075**	0.033	0.183
C. Perceived labor earnings			
Earnings compared to others at 25 (max10)	0.922***	0.301	6.755
Earnings compared to others at 35 (max10)	0.580***	0.223	7.817
D. Career aspration			
Standardized career aspiration pca	0.072	0.166	0.235
Number of observations	329		

Notes: Significance levels: * 10%, ** 5%, *** 1%. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Self-esteem index is constructed by applying a PCA using questions in columns (2)-(4). These questions are from Rosenberg's self-esteem scale (Rosenberg [1965]). Career aspiration index is constructed using a PCA using the questions from Gregor and O'Brien [2015] (Table A11).

Table 7: Effects of female teacher and gender composition for girls

	Dependent variable: Gender attitudes		
	(1)	(2)	(3)
STEM	0.261*	0.304**	0.226
	(0.139)	(0.133)	(0.141)
STEM × Female teacher	0.107		
	(0.196)		
STEM × 1(Num. of female students > Num. of male students)		-0.030	
		(0.212)	
STEM × 1(Above median avg. peers' pre-gender attitudes)			0.148
			(0.184)
Control group mean	.318	.318	.318
Observations	329	329	329

Notes: Significance levels: * 10%, ** 5%, *** 1%. Female teacher indicates the boot camps that had a female teacher as an instructor. Gender composition is obtained from the variations in the numbers of boys and girls in each boot camp. These regressions are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes. Center fixed effects are not included for estimating the propensity score because variations in teacher gender are highly correlated with the centers. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

Table 8: The effects of STEM bootcamps on behavior and long-term expectations for girls

Dependent variable	STEM	SE	Control mean
A. Weekly study hours			
Weekly hours of study	5.276***	1.947	11.479
Kiswahili/English weekly hours of study	1.174*	0.656	3.510
Math/Science/Computer weekly hours of study	2.202***	0.701	4.131
Geography/Civics/History weekly hours of study	1.900***	0.665	3.838
B. Educational goal			
Senior secondary or below	0.036	0.057	0.059
Technical/ vocational	-0.014	0.042	0.084
Higher education	-0.022	0.070	0.858
C. Preferred major in college			
Business	-0.101**	0.047	0.242
Engineering/Technology	0.189***	0.053	0.117
Other majors	-0.013	0.068	0.563
D. Expected occupation			
Clerical, sales and services	-0.058***	0.022	0.087
Education sector	0.006	0.057	0.112
Engineering/science/technology sector	0.067	0.063	0.324
Health and social/community work	0.028	0.041	0.091
Professional/ managerial (gov)	-0.030	0.043	0.261
Professional/ managerial (non-gov)	0.049	0.045	0.062
Agriculture/domestic service/manual job	-0.062***	0.023	0.062
E. Household outcomes			
Desired marriage age	0.089	0.472	26.458
Number of children wanted	-0.259*	0.141	3.588
Number of observations	329		

Notes: Significance levels: * 10%, ** 5%, *** 1%. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects.

Table 9: Heterogeneous effects by the days between STEM bootcamp and post-survey for girls

	Dependent variable: Gender attitudes
	(1)
STEM	0.338**
	(0.143)
STEM × 1(Above median days btw STEM camps and post surveys)	-0.077
	(0.183)
Control group mean	.318
Observations	329

Notes: Significance levels: * 10%, ** 5%, *** 1%. This regression result is estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

7 Appendix

Figure A1: Estimated propensity scores

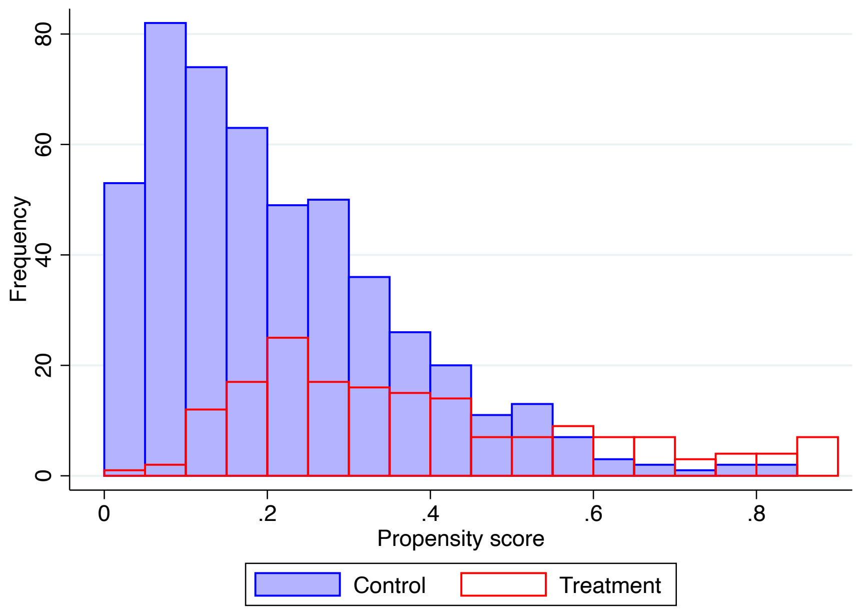


Table A1: Education system in Tanzania

Expected Age	7-13 years old	14-17 years old	18-19 years old	20-22 years old
School	Primary school	Secondary school (O-Level)	Secondary school (A-Level)	University
Grade	Standard 1- 7	Form 1 - 4	Form 5 - 6	Bachelor's degree

Notes: This table is for formal education and it does not include vocational track.

Table A2: Boot camp schedule

Time	Day 1	Day 2	Day 3	Day 4	Day 5
8:30-9:00	Team intro	Intro on lessons How computer works video	Recap session Echo show video	S4a - Switches and Sensors in Arduino	Motivation video Recaps Q/A
9:00-9:30	Growth mindset video	Commanding the robots by using Rogic	Victory game video Practice events	S4a -Buzzer in Arduino	Presentation preparation
9:30-10:00	Intro to block programming Real-life algorithm video	Test of the race bot Moving in different points on map	Victory game coding	LDR (light dep resistor) Ultrasonic sensor (wave)	Student group presentation preparation
10:30-11:00	Graph paper programming and debugging	Conclusion	Conclusion	Intro to mobile app App inventor features	Rachel server demo
11:00-11:30	BREAK	BREAK	BREAK	BREAK	BREAK
11:30-12:00	Intro lesson Change the world video	Intro lesson Line tracing bot video	Intro to the lesson on the maze	App inventor features Video	*all teachers and directors should be invited
12:00-12:30	Block coding: Star Wars coding	Coding of the robot Testing of the robot	Practice Exercise (hit the ball and say ouch!)	Understanding of blocks/ event handler/ procedure	Rachel demo continues
12:30-13:30	Amazing robots video	Disassembly of the robot	Maze game project		Preparation for closing ceremony
13:30-14:00	Conclusion	Conclusion	Extra projects; fruit catcher	Video Text-to-speech App	Interviews with Compassion staff and district school officials for internal documentation
14:00-15:00	LUNCH	LUNCH	LUNCH	LUNCH	LUNCH
15:00-15:30	Introducing lesson Arduino robot video	Review of programming Creating paper planes	Minecraft game Intro to s4a Electronic tools (s4a)	Text-to-speech App	Rachel server demo (optional); Closing ceremony Certificate awards
15:30-16:00	Assembly of racebot part 1	Introduction to scratch; example 1 (move Sprite 10-100 steps)	Code (s4a) reading session Making circuit LEDs	Accelerometer sensor app	Slide show - Practice group presentation to guests
16:00-16:30	Growth mindset video Assembly of racebot part 2	Coordinates in Scratch Getting Loopy Turn Sprite counterclockwise	Multiple LEDs Circuit with push button	Accelerometer sensor app	Disassembly of the robots and Arduino for the next camp Inventory checking Collection of handouts
16:30-17:00	Assembly of racebot part 3 Conclusion	Making Sprite have working animation Conclusion	Recap/review	Our Story App Install app on phone for demo	5 best students appreciation to the TAG center and DEO
17:00-18:00	Preparation for the next day	Group presentation preparation	Group presentation preparation and, sharing	Group sharing presentation preparations	

Table A3: Principal component loadings

	Principal component loadings
A woman's most important role is being a good home maker	0.4398
Husband and wife should both contribute to income	0.1425
Men make better political leaders than women do	0.4564
University is more important for a boy than for a girl	0.5111
Preschool child suffers with a working mother	0.0781
Women want a home and children	0.113
Men should take as much responsibility as women for home and children	0.0318
It is a problem if women have more income than husbands	0.3072
Having more than one wife is fine	0.4318
Wife must obey husband	-0.0341
Men make better business executives than women do	-0.1216

Notes: This table shows the principal component loadings for a PCA for baseline gender attitudes. We use answers for 11 questions (1-4 scale) on gender norms.

Table A4: Summary statistics

	STEM		Non-STEM		Diff		Diff	
	Mean	SD	Mean	SD	Coef.	SE	Coef.	SE
Age	14.724	1.582	14.955	1.527	-0.231*	0.138	0.008	0.110
Female	0.494	0.501	0.492	0.500	0.002	0.044	-0.022	0.047
Grade:								
Standard7	0.270	0.445	0.265	0.442	0.005	0.039	0.000	0.000
Form 1	0.322	0.469	0.362	0.481	-0.041	0.042	0.000	0.000
Form 2	0.299	0.459	0.209	0.407	0.090**	0.039	0.000**	0.000
Form 3	0.109	0.313	0.164	0.371	-0.055*	0.029	0.000*	0.000
Religion (base Islam):								
Christianity	0.746	0.435	0.603	0.486	0.143***	0.040	0.000***	0.000
Father in household	0.580	0.449	0.590	0.450	-0.010	0.040	-0.023	0.042
Mother in household	0.831	0.360	0.811	0.372	0.020	0.032	0.019	0.033
Father years of education	9.934	3.199	9.544	3.227	0.391	0.282	0.085	0.287
Mother years of education	8.987	2.857	8.571	3.408	0.416	0.265	0.201	0.278
Piped water source	0.777	0.413	0.706	0.452	0.071*	0.037	0.033	0.039
Electricity or solar energy source	0.715	0.450	0.673	0.466	0.042	0.040	0.000	0.041
Flush toilet	0.443	0.495	0.421	0.490	0.022	0.043	-0.033	0.045
Number of rooms in HHs	3.381	1.876	3.094	1.479	0.287*	0.157	0.254*	0.153
Computer use:								
Ever used a computer at school	0.160	0.348	0.177	0.363	-0.017	0.031	-0.015	0.033
Hours on math and science per week	4.758	5.156	6.456	6.971	-1.698***	0.501	-1.353**	0.533
Educational goal:								
Senior secondary or below	0.042	0.197	0.067	0.242	-0.025	0.018	-0.021	0.019
Technical/ vocational	0.101	0.297	0.115	0.309	-0.014	0.026	-0.016	0.028
Higher education	0.857	0.345	0.818	0.374	0.039	0.031	0.037	0.033
Expected occupation:								
Clerical, sales and services	0.018	0.131	0.017	0.126	0.001	0.011	0.001	0.011
Education sector	0.119	0.319	0.143	0.339	-0.024	0.029	-0.006	0.031
Engineering/science/technology sector	0.322	0.460	0.286	0.437	0.035	0.040	0.024	0.042
Health and social/community work	0.095	0.289	0.089	0.275	0.006	0.025	0.009	0.028
Professional/ managerial (government)	0.369	0.476	0.336	0.457	0.032	0.041	0.018	0.042
Professional/ managerial (non-government)	0.036	0.183	0.067	0.242	-0.031*	0.018	-0.033	0.020
Agriculture/domestic service/manual job	0.042	0.197	0.061	0.231	-0.019	0.018	-0.012	0.017
Desired marriage age (male)	28.193	4.600	27.366	4.513	0.827	0.566	0.757	0.597
Desired marriage age (female)	28.488	3.623	27.843	5.068	0.646	0.508	0.410	0.564
Standardized gender attitudes	0.406	1.010	0.000	1.000	0.406***	0.089	0.323***	0.089
Standardized self-esteem index	0.130	0.887	0.000	1.000	0.130	0.081	0.109	0.084
Big 5 personality:								
Extroversion	0.027	0.940	0.000	1.000	0.027	0.084	-0.003	0.088
Agreeableness	0.130	0.868	0.000	1.000	0.130	0.080	0.074	0.081
Conscientiousness	0.116	0.875	0.000	1.000	0.116	0.080	0.107	0.079
Neuroticism	0.090	0.872	0.000	1.000	0.090	0.080	0.076	0.082
Openness	0.062	0.932	0.000	1.000	0.062	0.084	0.030	0.080
Center fixed effects					No		Yes	
Control for selection criteria variables					No		Yes	
P-value for joint test of orthogonality					0.001		0.018	
Number of observations	174		494		668		668	

Notes: Significance levels: * 10%, ** 5%, *** 1%. The samples include all the students who we observe both in the baseline and endline surveys. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Self-esteem index is constructed using a PCA using three questions from the Rosenberg's self-esteem scale. Selection criteria variables include class year (grade) and religion. For indexes and personalities, mean and standard deviation are normalized to 0 and 1.

Table A5: Attrition

Dependent variable:	Attrited (= 1)
STEM	-0.192***
	(0.024)
Age	-0.012
	(0.011)
Female	0.799
	(0.736)
Grade: Standard7	0.055
	(0.052)
Grade: Form 1	-0.016
	(0.042)
Grade: Form 2	0.045
	(0.041)
Religion: Christianity	0.024
	(0.027)
Father in household	-0.008
	(0.030)
Mother in household	-0.035
	(0.037)
Piped water source	-0.022
	(0.032)
Electricity or solar energy source	0.040
	(0.031)
Flush toilet	-0.029
	(0.029)
Number of rooms in HHs	0.006
	(0.009)
Computer use: No computer or never use at school	-0.014
	(0.037)
Hours on math and science per week	-0.001
	(0.002)
Personality: Extroversion	-0.000
	(0.017)
Personality: Agreeableness	-0.038**
	(0.018)
Personality: Conscientiousness	0.006
	(0.016)
Personality: Neuroticism	0.003
	(0.016)
Personality: Openness	-0.017
	(0.017)
Educational goal: Senior secondary or below	0.160***
	(0.058)
Educational goal: Technical/vocational	-0.025
	(0.040)
Desired marriage age (male)	0.002
	(0.004)
Desired marriage age (female)	-0.004
	(0.004)
Expected occupation: Clerical, sales and services	0.012
	(0.137)
Expected occupation: Education sector	-0.018
	(0.077)
Expected occupation: Engineering/science/technology sector	-0.134*
	(0.071)
Expected occupation: Health and social/community work	-0.043
	(0.079)
Expected occupation: Professional/ managerial (government)	-0.151**
	(0.070)
Expected occupation: Professional/ managerial (non-government)	0.029
	(0.090)
Standardized gender attitudes	-0.001
	(0.014)
Standardized self-esteem index	0.033**
	(0.013)
Observations	829

Notes: Significance levels: * 10%, ** 5%, *** 1%. This sample includes all the students who we observe in the baseline survey.

Table A6: The effects of the STEM bootcamp on gender attitudes

Dependent variable:	Gender attitudes simple average index			Gender attitudes Anderson index		
	(1) OLS	(2) OLS	(3) Inverse probability weighting	(4) OLS	(5) OLS	(6) Inverse probability weighting
STEM	0.215**	0.177**	0.135*	0.173**	0.154*	0.125*
	(0.086)	(0.089)	(0.081)	(0.072)	(0.079)	(0.071)
Control baseline gender attitudes	Yes	Yes	IPW	Yes	Yes	IPW
Control covariates	No	Yes	IPW	No	Yes	IPW
Control group mean	0	0	0	0	0	0
R-squared	0.123	0.208		0.125	0.186	
Observations	668	668	668	668	668	668

Notes: Significance levels: * 10%, ** 5%, *** 1%. Covariate variables include all variables available in the baseline survey and squared and cubic terms for continuous variables. Center fixed effects are included in all specifications. Propensity score for inverse probability weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. For continuous variables, squared and cubic terms are also included. Gender attitudes in column (1)-(3) are constructed by simply taking averages of the answers (1-4 scale). Gender attitudes in column (4)-(6) are constructed following Anderson [2008] and Dhar et al. [2020]. All the outcomes are normalized so that the control group baseline mean is 0 and standard deviation is 1.

Table A7: Degree of selection on unobservables

	(1)		(2)		Degree of selection relative to (1)
	Coef.	SD	Coef.	SD	δ
A. Gender attitudes					
Standardized gender attitudes index	0.189**	0.079	0.183**	0.084	3.69
B. Self-esteem					
Standardized self-esteem index	0.311***	0.088	0.354***	0.092	-26.44
C. Favorite subjects					
1st favorite subject-math and science	0.151***	0.041	0.131***	0.044	3.09
2nd favorite subject-math and science	0.166***	0.043	0.178***	0.046	10.43
3rd favorite subject-math and science	-0.074**	0.030	-0.053	0.032	2.5
D. Career aspiration					
Standardized career aspiration pca	0.081	0.135	0.134	0.144	-3.78
E. Earnings compared to others (1-10 scale)					
Earnings compared to others at 25 (max10)	0.417**	0.189	0.394*	0.210	1.36
Earnings compared to others at 35 (max10)	0.331**	0.154	0.296*	0.171	.79
F. Weekly study hours					
Weekly hours of study	3.697***	0.969	3.511***	0.965	1.62
Kiswahili/English weekly hours of study	1.007***	0.319	0.979***	0.328	1.5
Math/Science/Computer weekly hours of study	1.592***	0.390	1.479***	0.381	1.91
Geography/Civics/History weekly hours of study	1.098***	0.323	1.053***	0.329	1.56
G. Educational goal					
Senior secondary or below	-0.027	0.022	-0.039*	0.023	-4.02
Technical/ vocational	-0.051**	0.023	-0.040	0.026	1.25
Higher education	0.078**	0.031	0.079**	0.033	4.9
H. Preferred major in college					
Business	-0.113***	0.033	-0.102***	0.035	3.6
Engineering/Technology	0.193***	0.041	0.208***	0.043	9.32
Other majors	-0.033	0.046	-0.061	0.047	-3.11
I. Expected occupation					
Clerical, sales and services	-0.051***	0.018	-0.060***	0.021	14.44
Education sector	0.009	0.029	0.019	0.031	-2.55
Engineering/science/technology sector	0.145***	0.044	0.146***	0.046	2.98
Health and social/community work	0.001	0.028	0.001	0.031	-1.16
Professional/ managerial (gov)	-0.043	0.033	-0.059	0.038	-1.81
Professional/ managerial (non-gov)	-0.007	0.021	0.013	0.024	-.83
Agriculture/domestic service/manual job	-0.054***	0.020	-0.060***	0.022	14.88
J. Desired marriage age					
Desired marriage age	0.510	0.418	0.257	0.411	1.1
K. Desired number of children					
Desired number of children	-0.329**	0.144	-0.331**	0.141	8.21
Control baseline outcomes	Yes (if exists)		Yes (if exists)		
Control covariates	No		Yes		
Number of observations	668		668		

Notes: Significance levels: * 10%, ** 5%, *** 1%. The degree of selection δ values are constructed using the method in Oster [2019]. These values indicate the degree of selection on unobservables relative to observables needed to explain away the estimated impacts. As suggested by Oster [2019], we use $R_{max} = 1.3 \times R^2$ where R^2 is obtained from the saturated model. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Self-esteem index is constructed using a PCA using three questions from the Rosenberg's self-esteem scale. Career aspiration index is constructed using a PCA using the questions from Gregor and O'Brien [2015].

Table A8: Attrition bounds

	Upper bounds		Lower bounds	
	Coef.	SD	Coef.	SD
A. Gender attitudes				
Standardized gender attitudes index	0.266***	0.079	0.157**	0.078
B. Self-esteem				
Standardized self-esteem index	0.389***	0.088	0.246***	0.088
C. Favorite subjects				
1st favorite subject-math and science	0.181***	0.040	0.093**	0.040
2nd favorite subject-math and science	0.209***	0.042	0.120***	0.043
3rd favorite subject-math and science	-0.034	0.028	-0.034	0.028
D. Career aspiration				
Standardized career aspiration pca	0.122	0.128	0.005	0.132
E. Earnings compared to others (1-10 scale)				
Earnings compared to others at 25 (max10)	0.497***	0.188	0.142	0.195
Earnings compared to others at 35 (max10)	0.389**	0.153	0.034	0.165
F. Weekly study hours				
Weekly hours of study	4.265***	0.927	3.200***	0.922
Kiswahili/English weekly hours of study	1.212***	0.310	0.857***	0.309
Math/Science/Computer weekly hours of study	1.736***	0.364	1.381***	0.362
Geography/Civics/History weekly hours of study	1.317***	0.310	0.961***	0.308
G. Educational goal				
Senior secondary or below	-0.023	0.022	-0.023	0.022
Technical/ vocational	-0.025	0.023	-0.025	0.023
Higher education	0.048	0.030	0.048	0.030
H. Preferred major in college				
Business	-0.068**	0.031	-0.080**	0.031
Engineering/Technology	0.214***	0.040	0.214***	0.040
Other majors	-0.018	0.042	-0.107**	0.044
I. Expected occupation				
Clerical, sales and services	-0.046**	0.018	-0.046**	0.018
Education sector	0.028	0.028	0.028	0.028
Engineering/science/technology sector	0.169***	0.042	0.081*	0.043
Health and social/community work	0.016	0.027	0.016	0.027
Professional/ managerial (gov)	-0.047	0.034	-0.047	0.034
Professional/ managerial (non-gov)	0.007	0.020	0.007	0.020
Agriculture/domestic service/manual job	-0.038*	0.020	-0.038*	0.020
J. Desired marriage age				
Desired marriage age	0.558	0.371	0.114	0.385
K. Desired number of children				
Desired number of children	-0.204	0.124	-0.292**	0.124
Number of observations	829		829	

Notes: Significance levels: * 10%, ** 5%, *** 1%. Missing observations are set equal to the 25th percentile value for each outcome to obtain upper bounds (negative selection). Missing observations are set to the 75th percentile value to construct lower bounds (positive selection). Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Self-esteem index is constructed using a PCA using three questions from the Rosenberg's self-esteem scale. Career aspiration index is constructed using a PCA using the questions from Gregor and O'Brien [2015].

Table A9: Summary statistics by gender

	Males						Females					
	STEM		Non-STEM		Diff (w/ center FE)		STEM		Non-STEM		Diff (w/ center FE)	
	Mean	SD	Mean	SD	Coef.	SE	Mean	SD	Mean	SD	Coef.	SE
Age	15.131	1.567	15.112	1.548	0.042	0.233	14.892	1.637	14.747	1.512	0.116	0.226
Grade:												
Standard7	0.177	0.384	0.228	0.421	-0.056	0.049	0.243	0.432	0.290	0.454	-0.026	0.058
Form 1	0.421	0.497	0.359	0.481	0.062	0.069	0.321	0.470	0.325	0.469	-0.009	0.073
Form 2	0.258	0.440	0.249	0.433	0.015	0.066	0.271	0.447	0.242	0.429	0.033	0.065
Form 3	0.145	0.354	0.163	0.370	-0.021	0.049	0.164	0.373	0.144	0.352	0.002	0.055
Religion (base Islam):												
Christianity	0.691	0.460	0.633	0.479	0.052	0.067	0.667	0.474	0.654	0.472	0.016	0.073
Father in household	0.539	0.454	0.595	0.443	-0.058	0.061	0.617	0.449	0.622	0.449	-0.013	0.069
Mother in household	0.829	0.363	0.833	0.356	0.003	0.052	0.825	0.369	0.812	0.372	0.005	0.054
Father years of education	9.784	2.944	9.442	3.130	0.284	0.368	10.099	3.357	9.919	3.494	0.116	0.485
Mother years of education	8.916	3.032	8.576	3.220	0.230	0.396	8.957	3.000	8.908	3.639	0.027	0.430
Piped water source	0.725	0.446	0.733	0.441	-0.017	0.067	0.751	0.427	0.716	0.448	0.028	0.064
Electricity or solar energy source	0.639	0.481	0.656	0.472	-0.026	0.073	0.745	0.433	0.710	0.451	0.026	0.057
Flush toilet	0.422	0.497	0.401	0.486	0.015	0.067	0.479	0.493	0.460	0.496	0.015	0.076
Number of rooms in HHs	3.275	1.825	3.197	1.497	0.075	0.231	3.053	1.475	3.102	1.433	-0.053	0.180
Computer use:												
Ever used a computer at school	0.160	0.342	0.147	0.336	0.017	0.051	0.250	0.420	0.191	0.373	0.049	0.059
Hours on math and science per week	5.645	6.550	5.921	6.830	-0.216	0.935	5.360	4.987	6.020	6.205	-0.625	0.796
Educational goal:												
Senior secondary or below	0.064	0.235	0.079	0.262	-0.016	0.037	0.025	0.156	0.038	0.186	-0.013	0.018
Technical/ vocational	0.128	0.324	0.139	0.334	-0.010	0.057	0.128	0.334	0.089	0.280	0.030	0.052
Higher education	0.809	0.380	0.782	0.400	0.026	0.064	0.847	0.360	0.873	0.326	-0.017	0.054
Expected occupation:												
Clerical, sales and services	0.017	0.124	0.019	0.134	-0.003	0.015	0.008	0.089	0.015	0.116	-0.008	0.011
Education sector	0.137	0.335	0.130	0.328	0.006	0.047	0.120	0.325	0.135	0.329	-0.010	0.046
Engineering/science/technology sector	0.289	0.438	0.299	0.446	-0.008	0.060	0.317	0.467	0.308	0.446	0.001	0.072
Health and social/community work	0.081	0.264	0.079	0.263	0.002	0.035	0.102	0.304	0.101	0.291	0.011	0.043
Professional/ managerial (government)	0.381	0.471	0.321	0.455	0.058	0.063	0.367	0.483	0.371	0.467	-0.001	0.068
Professional/ managerial (non-government)	0.040	0.188	0.073	0.252	-0.031	0.025	0.038	0.192	0.042	0.193	-0.007	0.027
Agriculture/domestic service/manual job	0.054	0.219	0.078	0.261	-0.023	0.028	0.049	0.216	0.028	0.159	0.015	0.040
Desired marriage age	27.493	4.573	27.408	4.381	0.014	0.636	28.317	3.472	27.947	4.821	0.343	0.541
Standardized gender attitudes	-0.041	0.990	-0.211	1.006	0.176	0.133	0.494	1.009	0.464	0.867	0.035	0.170
Standardized self-esteem index	0.117	0.930	0.093	0.902	0.018	0.134	0.072	0.996	0.029	1.022	0.033	0.155
Big 5 personality:												
Extroversion	0.062	0.938	-0.061	1.076	0.094	0.138	0.125	0.734	0.126	0.818	0.002	0.091
Agreeableness	0.160	0.808	0.017	1.007	0.113	0.121	0.139	0.825	0.084	0.872	0.047	0.101
Conscientiousness	0.162	0.740	-0.021	1.051	0.161	0.106	0.169	0.773	0.137	0.784	0.031	0.093
Neuroticism	0.083	0.822	-0.067	1.079	0.129	0.116	0.094	0.867	0.147	0.786	-0.067	0.130
Openness	0.129	0.788	0.009	1.005	0.102	0.101	0.113	0.869	0.081	0.882	0.029	0.118
P-value for joint test of orthogonality					0.984						1.000	
Number of observations	88		251				86		243			

Notes: Significance levels: * 10%, ** 5%, *** 1%. The samples include all the students who we observe both in the baseline and endline surveys. The results in the last four columns are obtained from the regressions using an inverse probability weighting. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Self-esteem index is constructed using a PCA using three questions from the Rosenberg's self-esteem scale.

Table A10: The effects of the STEM bootcamp on gender attitudes for girls (propensity score matching)

	Dependent variable: Gender attitudes		
	(1)	(2)	(3)
STEM	0.288**	0.315**	0.277**
	(0.129)	(0.145)	(0.140)
Nearest neighbor	1	2	3
Control group mean	.318	.318	.318
Observations	329	329	329

Notes: Significance levels: * 10%, ** 5%, *** 1%. Propensity score is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers (1-4 scale) for 11 questions on gender norms. Unweighted control mean and standard deviation are normalized to 0 and 1.

Table A11: Questionnaires for career aspirations

1. I hope to become a leader in what I do as an adult.
2. I do not plan to devote energy to getting promoted to a leadership position in the organization or business in which I am working.
3. I want to be among the very best in what I do in my career.
4. Becoming a leader in my job is not at all important to me.
5. When I am established in my career, I would like to manage other workers.
6. I plan to reach the highest level of education in my field.
7. I want to have responsibility for the future direction of my organization or business.
8. I want my work to have a lasting impact on my field.
9. I aspire to have my contributions at work recognized by my boss.
10. I will pursue additional training in my occupational area of interest.
11. I will always be knowledgeable about recent advances in my field.
12. Attaining leadership status in my career is not that important to me.
13. Being outstanding at what I do at work is very important to me.
14. I know I will work to remain current regarding knowledge about my work.
15. I hope to move up to a leadership position in my organization or business.
16. I will attend conferences annually to advance my knowledge.
17. I know that I will be recognized for my accomplishments in what I do in my career.
18. Even if not required, I would take continuing education courses to become more knowledgeable.
19. I would pursue an advanced education program to gain specialized knowledge in what I do in my career.
20. Achieving in my career is not at all important to me.
21. I plan to obtain many promotions in my organization or business.
22. Being one of the best in my field is not important to me.
23. Every year, I will prioritize involvement in continuing education to improve my ability for my job
24. I plan to rise to the top leadership position of my organization or business.

Notes: These questions are from Gregor and O'Brien [2015]. Students were asked to choose a number between 0 (not at all true of me) and 4 (very true of me).

Table A12: Variations in teacher gender and sex ratios during STEM boot camps

Site	Week	Teacher gender composition	Ratio of male to female students
AICT Kitete	1	(Male, Male)	1.22
AICT Kitete	2	(Male, Male)	1.00
AICT Kitete	3	(Male, Male)	1.43
Monravian Chemchem	1	(Male, Male)	0.71
Monravian Chemchem	2	(Male, Male)	0.60
Monravian Chemchem	3	(Male, Male)	1.60
Angelican Isevya	1	(Male, Female)	1.64
Angelican Isevya	2	(Male, Female)	0.44
Angelican Isevya	3	(Male, Female)	1.00

Notes: This table shows variations in teacher gender composition and sex ratios of students during the STEM boot camps in each week and site.

Table A13: Heterogeneous effects by social desirability index for boys

	Dependent variable: Gender attitudes
	(1)
STEM	-0.053
	(0.149)
STEM $\times$ Above median social desirability index	0.158
	(0.209)
Above median social desirability index	0.007
	(0.178)
Control group mean	-.308
Observations	339

Notes: Significance levels: * 10%, ** 5%, *** 1%. Social desirability index is constructed by applying a PCA using 3 baseline questions ("I feel little concern for others", "I insult people" and "I take time out for others") in big 5 personalities. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

Table A14: The effects of STEM boot camp for boys

Dependent variable	STEM	SE	Control mean
A. Self-esteem			
Standardized self-esteem index	0.425***	0.099	0.058
B. Interests in STEM subjects			
1st favorite subject-math and science	0.163***	0.047	0.611
2nd favorite subject-math and science	0.200***	0.059	0.439
3rd favorite subject-math and science	-0.026	0.053	0.143
C. Perceived labor earnings			
Earnings compared to others at 25 (max10)	0.146	0.239	6.593
Earnings compared to others at 35 (max10)	0.301	0.208	7.521
D. Career aspration			
Standardized career aspiration pca	0.231	0.192	-0.227
Number of observations	339		

Notes: Significance levels: * 10%, ** 5%, *** 1%. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects. Self-esteem index is constructed by applying a PCA using questions in columns (2)-(4). These questions are from Rosenberg's self-esteem scale (Rosenberg [1965]). Career aspiration index is constructed using a PCA using the questions from Gregor and O'Brien [2015] (Table A11).

Table A15: Effects of female teacher and gender composition for boys

	Dependent variable: Gender attitudes		
	(1)	(2)	(3)
STEM	0.136	0.065	-0.136
	(0.155)	(0.158)	(0.178)
STEM × Female teacher	-0.234		
	(0.251)		
STEM × 1(Num. of female students > Num. of male students)		-0.130	
		(0.280)	
STEM × 1(Above median avg. peers' pre-gender attitudes)			0.394*
			(0.232)
Control group mean	-.308	-.308	-.308
Observations	339	339	339

Notes: Significance levels: * 10%, ** 5%, *** 1%. Female teacher indicates the boot camps that had a female teacher as an instructor. Gender composition is obtained from the variations in the numbers of boys and girls in each boot camp. These regressions are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes. Center fixed effects are not included for estimating the propensity score because variations in teacher gender are highly correlated with the centers. Gender attitudes are constructed by applying a PCA to answers for 11 questions (1-4 scale) on gender norms.

Table A16: The effects of STEM bootcamps on behavior and long-term expectations for boys

Dependent variable	STEM	SE	Control mean
A. Weekly study hours			
Weekly hours of study	3.916***	1.043	11.873
Kiswahili/English weekly hours of study	1.076***	0.312	3.654
Math/Science/Computer weekly hours of study	1.984***	0.487	4.171
Geography/Civics/History weekly hours of study	0.856**	0.346	4.048
B. Educational goal			
Senior secondary or below	-0.063***	0.024	0.093
Technical/ vocational	-0.056*	0.033	0.146
Higher education	0.119***	0.039	0.760
C. Preferred major in college			
Business	-0.114***	0.041	0.282
Engineering/Technology	0.163***	0.055	0.176
Other majors	-0.027	0.064	0.471
D. Expected occupation			
Clerical, sales and services	-0.063***	0.018	0.073
Education sector	0.012	0.037	0.109
Engineering/science/technology sector	0.219***	0.061	0.287
Health and social/community work	-0.008	0.034	0.081
Professional/ managerial (gov)	-0.109***	0.040	0.243
Professional/ managerial (non-gov)	-0.014	0.028	0.073
Agriculture/domestic service/manual job	-0.037	0.036	0.134
E. Household outcomes			
Desired marriage age	0.880*	0.497	27.429
Number of children wanted	-0.357**	0.162	4.041
Number of observations	339		

Notes: Significance levels: * 10%, ** 5%, *** 1%. These regression results are estimated using an inverse probability weighting. Propensity score for the weighting is estimated using all variables that are available in the baseline survey including the baseline gender attitudes and center fixed effects.