Reshaping Adolescents' Gender Attitudes:
Evidence from a School-Based Experiment in India

Diva Dhar   Tarun Jain   Seema Jayachandran[*]

November 2, 2021

Abstract

This paper evaluates an intervention in India that engaged adolescent girls and boys in classroom discussions about gender equality for two years, aiming to reduce their support for societal norms that restrict women's and girls' opportunities. Using a randomized controlled trial, we find that the program made attitudes more supportive of gender equality by 0.18 standard deviations, or, equivalently, converted 16% of regressive attitudes. When we resurveyed study participants two years after the intervention had ended, the effects had persisted. The program also led to more gender-equal self-reported behavior, and we find weak evidence that it affected two revealed-preference measures.

JEL Codes: J12, J13, J16, O12.

Gender inequality exists in every society, but it is especially acute in many developing countries. Compared to men and boys, women and girls have fewer educational opportunities, less autonomy in decisions about their marriage and fertility, and more restrictions on their labor market participation and even their physical mobility and friendships (Duflo, 2012; Jayachandran, 2015). Many of these gender gaps have not narrowed in recent decades despite economic progress. Cultural norms often underpin these disparities, and economic development alone is unlikely to eliminate them (Alesina et al., 2013; Jayachandran, 2021).

Various policies might help foster greater gender equality in the face of restrictive gender norms. These include laws that guarantee equal rights, subsidies to encourage investment in girls, and programs that impart skills or target resources to women. A large literature has studied these approaches.

This paper focuses on a less common—and less commonly studied—approach: trying to directly change people's gender attitudes. We study an intervention that used discussion and persuasion to reduce participants' support for restrictive gender norms and to increase the value they place on equality. That is, it aimed to change their preferences.

Gender attitudes, even those rooted in centuries-old cultural norms, are amenable to change. For example, reserving seats for female politicians has helped curtail negative stereotypes about women as leaders in India (Beaman et al., 2009), and television shows have changed fertility preferences in multiple settings (Jensen and Oster, 2009; La Ferrara et al., 2012). The distinctiveness of the intervention we study is that reshaping gender attitudes was its primary goal.

The intervention, which we evaluate through a randomized controlled trial, was implemented in secondary schools in the state of Haryana, India. It engaged seventh to tenth graders in classroom discussions about gender equality, with a 45-minute session held every three weeks for two and a half school years. The sessions taught facts and endorsed gender equality and, as importantly, prompted students to reflect on their own and society's views. Discussion topics included gender stereotypes, gender roles at home, girls' education, women's employment outside the home, and harassment. Some sessions taught communication skills to help students convey their views to others so that they can, for example, persuade their parents to permit them to marry at a later age. The program's messaging combined a human-rights case for gender equity with pragmatic reasons to value women, such as their economic contributions.

Breakthrough, a non-profit organization with extensive experience in gender-equality

1

programming, designed and implemented the intervention. The Government of Haryana allowed Breakthrough to conduct the classes in schools as part of the regular school day. The government was interested in eroding some of the existing gender norms, and schools offer governments a powerful platform to shape the next generation's views. Secondary school students were specifically targeted because adolescence is a critical time in the development of morality and identity formation, with adolescents being young enough to still have malleable attitudes but mature enough to reflect on complex moral questions (Kohlberg, 1976; Markus and Nurius, 1986).[1]

Gender inequality is rife in India. While boys and girls start secondary school at the same rate, only 0.80 girls enroll in tertiary schooling for every boy (World Bank, 2011). Early marriage is common, and many women have limited agency (Kishor and Gupta, 2004). India's female labor force participation rate is among the lowest in the world (Field et al., 2010; Klasen and Pieters, 2015; Afridi et al., 2018). It also has one of the most male-skewed sex ratios worldwide, due to the widespread practice of sex-selective abortion (Sen, 1990). Among children age 0 to 6 years, there are 1.09 boys per girl; Haryana's sex ratio of 1.20 is the most male-skewed among Indian states (Jha et al., 2006; Govt. of India, 2011).

Our study encompasses 314 government secondary schools across four districts in Haryana. We collected data from roughly 14,000 students, both boys and girls. We analyze the intervention's effects a few months after the program ended and again two years later when the students were, on average, 17 years old.

The main outcome we study is participants' gender attitudes, that is, their views about what is right and wrong or desirable and undesirable, such as whether it is wrong for women to work outside the home and whether it would be good to have more women in politics. The second outcome is educational and career aspirations; for this outcome, we hypothesized that the intervention would only influence girls. The third outcome is self-reported behaviors influenced by gender norms. Only when the participants become adults can we assess impacts on major outcomes that they have control over, such as employment and child-bearing. Nonetheless, examining day-to-day behavior in adolescence can provide insights on whether reshaping gender attitudes translates into behavior change. We focus on behaviors that adolescents likely have some say over, such as chores done at home and interaction with opposite-gender peers. One reason why attitude change might be insufficient for behavior change is people's desire to conform to social norms. A boy who believes that he should help

---

[1]There were also practical reasons for targeting adolescents rather than younger students. The government was less comfortable with younger children discussing possibly sensitive topics like sexual harassment.

out with chores might worry about the social sanctions he would face if he did so. For this reason, we examine perceptions of social norms as a secondary outcome.

We find that the intervention made gender attitudes more progressive by 0.18 standard deviations in the short run (three and a half months after the program ended, or three years after baseline). The measure of attitudes is an index that aggregates several survey responses pertaining to support for gender equality. The effect size is equivalent to newfound support for gender equality in 16% of the cases where a student initially held a gender-regressive view.

What is especially striking is that these effects persisted. We continue to find a large effect on attitudes—0.16 standard deviations—in the medium run (two years after the program ended, or five years after baseline).

The program also influenced participants' self-reported behavior. An index of gender-equal behaviors increased by 0.20 to 0.23 standard deviations, in both the short run and medium run. In addition, we added two objective measures of behavior to the second endline. The first was a revealed preference measure of girls' educational intentions, namely whether they submitted a college scholarship application. The second measure tracked boys' and girls' likelihood of signing a public petition to end the dowry system. We find weak evidence of a treatment effect in the hypothesized direction for both behaviors. Finally, we find no evidence that the intervention increased girls' stated educational and career aspirations, which were quite high to begin with.

We investigate how the treatment effects vary based on two pre-specified characteristics, student gender and parents' gender attitudes. We find no evidence of heterogeneity based on parents' attitudes, but we find important differences in treatment effects between boys and girls. At the first endline, the effect size on attitudes is somewhat larger for boys than girls, but we cannot statistically reject that the effect size is identical for the two groups. By the second endline, the effect on attitudes is significantly larger for boys. In addition, behavior change is significantly more pronounced among boys in both the short and medium run. For example, boys report doing more chores, but girls do not report doing fewer, and only boys report an increase in how much they encourage their older sisters to pursue a college education. With chores, the smaller effect for girls can be explained by the asymmetry in the outcome. Greater gender equality maps to girls seeking a lower burden of chores, which others might not grant them, while boys can voluntarily help out more. But the fact that we also see heterogeneity by gender in support given to sisters suggests a broader phenomenon

3

of girls facing more constraints on their behavior. This pattern highlights that, because behavior change requires not just the desire but also the ability to act differently, the very fact of boys' and men's greater power in society makes it important to include them in interventions aimed at increasing girls' and women's power.

A key concern with self-reported outcomes such as gender attitudes is the possibility of social desirability bias. The specific concern in our experiment is that there might be more social desirability bias in the treatment group, i.e., there could be experimenter demand effects. Participants in a program that explicitly tried to influence their support for gender equality might disingenuously express more gender-progressive views to present themselves in a good light to the surveyors. The persistence of the effects two years after the program ended provides some reassurance that it genuinely changed participants' views; it seems likely that experimenter demand effects would fade out with time. However, to address this concern more rigorously, we use the Marlowe-Crowne social desirability scale, which is a survey module developed by social psychologists to measure a person's propensity to give socially desirable answers (Crowne and Marlowe, 1960). The module, which we included in the baseline survey, asks respondents if they have several too-good-to-be-true traits such as never being jealous of another person's good fortune and always being a good listener; those who report more of these traits are scored as having a higher propensity to give socially desirable answers.

We find that respondents with a high propensity for social desirability bias express more support for gender equality overall for the sample, but importantly, this pattern is not truer for the treatment group than the control group. In other words, the positive treatment effects on self-reported attitudes and behavior are similar in magnitude for respondents with a low versus high propensity for social desirability bias. We view this analysis as an important check on the validity of our results, and think that our approach could be useful in a wide array of studies in which experimenter demand is a concern. The method allows one to test for bias for any and all outcomes, so it complements techniques such as list experiments and revealed-preference measures, which often must focus on a narrower set of outcomes for logistical reasons.

Our study contributes to the literature on endogenous preferences, specifically on the formation of gender-related preferences.[2] Besides political quotas (Beaman et al., 2009)

---

[2]Recent work has also studied the formation of preferences toward different castes and toward children from poorer families in India (Lowe, 2021; Rao, 2019), Muslims in the United Kingdom (Alrababa'h et al., 2019), and racial minorities in the US (Carrell et al., 2019).

and television (Jensen and Oster, 2009; La Ferrara et al., 2012), other factors that have been shown to positively influence gender attitudes include mothers' employment (Fernandez et al., 2004), having daughters or sisters (Washington, 2008; Healy and Malhotra, 2013), serving with women in the military (Dahl et al., 2021), having teachers who hold weaker gender stereotypes (Carlana, 2019), and having female role models (Porter and Serra, 2019).

Unlike most of the studies above on endogenous gender preferences, our research examines an intervention that intentionally changed preferences. It thus also sits within the literature on persuasion, or communication expressly designed to change preferences or beliefs (DellaVigna and Gentzkow, 2010). Much of the economics literature on persuasion focuses on ways to influence consumer or political preferences. Closer to our work are studies across the social sciences on attitude change related to intimate partner violence (Gupta et al., 2013; Abramsky et al., 2014; Pulerwitz et al., 2015; Green et al., 2020), racial minorities (Donovan and Leivers, 1993), immigrants (Hopkins et al., 2019; Grigorieff et al., 2020), and women in STEM (Moss-Racusin et al., 2018), as well as studies that shift people's perceptions of social norms about gender or about ethnic discrimination and violence (Bursztyn et al., 2020; Aloud et al., 2020; Paluck, 2009). Our study is also related to Cantoni et al. (2017), which finds that Chinese students taught with textbooks designed to convey pro-Communist messages express more pro-government views and skepticism of free markets as adults.

We also add to the rapidly growing literature on educational/training interventions aimed at increasing girls' and women's agency and opportunities in developing countries. Related work includes Bandiera et al. (2020) on female empowerment and livelihood training in Uganda, Buchmann et al. (2018) on empowerment training and financial incentives to delay marriage in Bangladesh, Ashraf et al. (2020) on negotiation skills training for girls in Zambia, Edmonds et al. (2020) on life skills training for girls in India, and McKelway (2020) on self-efficacy training in India. While most of these interventions primarily impart human capital to women and girls—either traditional skills or positive psychological traits—the central (though not exclusive) aim of the intervention we evaluate is to influence participants' preferences, specifically their attitudes about traditional gender roles. Our study is also relatively unusual in the literature on women's empowerment in that it focuses on both boys and girls.

I. Description of the intervention

This project emerged from the Government of Haryana's interest in testing policies to narrow gender gaps in its society. Breakthrough, a human rights organization specializing in social change campaigns related to gender, designed an intervention aimed at changing adolescent boys' and girls' views about gender norms, and implemented it in government schools with the state government's permission.

The participants in the program were the cohorts in grades 7 and 8 in the academic year 2014-15 when the program launched. It ran from April 2014 to October 2016, so one cohort participated in the program in grades 7, 8 and half of 9, and the other in grades 8, 9 and half of 10. Grades 7 to 10 have high enrollment and low dropout in Haryana, so the program could reach a large share of the underlying age cohorts and have limited attrition due to school dropout (significant dropout occurs after grade 10) (DISE, 2011).

The objective of the program, which was named *Taaron ki Toli*, or Legion of Stars, was to create awareness of gender-based discrimination, change dominant gendered perceptions, promote gender-equitable attitudes, raise girls' aspirations, and provide tools to participants to translate attitude change and greater aspirations into behavior change. The program aimed to ultimately influence a wide range of behaviors related to female education, mobility, work, marriage, and fertility, for both female participants and male participants' female family members (e.g., their future wives).

The program emphasized both human rights and pragmatic (i.e., instrumental) reasons for giving girls and women more opportunities. For example, it conveyed that equal opportunity for education is a universal human right. The hypothesis is that this rights-based message would increase how much participants value girls having access to higher education; a boy would get disutility from seeing his sister denied the same chance to attend college that he has. As an example of a pragmatic argument for girls' education, the intervention informed participants that outcomes for children improve when their mother is more educated. The hypothesis is that this information causes updating of beliefs; girls want to stay in school longer, and both boys and girls will want to educate their daughters down the road.

To ensure that the intervention would be widely accepted, Breakthrough engaged with multiple stakeholders at the state, district, and sub-district levels, gathering input from education officials, school principals, and teachers as they developed the program. This helped them design a program that might later be integrated into the standard school curriculum.[3]

---

[3]Scale-up options include hiring special-purpose teachers to lead the sessions, each of whom covers multiple

The program featured classroom sessions, each 45 minutes long, led by a Breakthrough facilitator. The regular teacher was welcome to stay for the sessions or leave the classroom and have a break. There were a total of 27 sessions spread over two and half years.[4] Breakthrough hired 15 facilitators, 13 of whom were male, to cover the 150 treatment schools.[5] The facilitator visited each school roughly once every three weeks. Other elements of the program included a one-time training for one teacher per school, optional youth clubs, and school-wide activities such as street theater performances held about once a year.[6]

The classroom sessions were discussion-based, with more student participation than is typical in secondary schools classes. Perhaps for this reason, and also because having a dynamic teaching style was one of the hiring criteria for facilitators, student reaction to the program was generally very positive. To complement the in-class material and encourage further reflection, the facilitators assigned some homework assignments such as to write stories and record observations, and they encouraged students to talk to their family members about what they were learning. The sessions did not displace one specific subject like math or history; they crowded out a roughly even mix of material in other subjects. While the program could have harmed other learning by displacing instructional time, the discuss-and-debate style could have strengthened students' critical thinking and speaking skills, conferring benefits beyond the gender focus.

Discussion topics for the sessions included gender identity, values, aspirations, gender

schools; having regular school teachers deliver the lessons; incorporating some of the content into textbooks; or continuing to have NGOs implement the program. The Government of Punjab (India) announced in March 2021 that it planned to roll out the program in all of its upper primary schools in late 2021, to be taught by regular social studies and moral science teachers (Menon, 2021).

[4]The total dosage was 20 hours, or about 23 hours adding in the school-wide assemblies. As comparisons, the negotiation program for girls in Zambia studied by Ashraf et al. (2020) consisted of six two-hour sessions (12 hours total); the safe space groups in Bangladesh evaluated by Buchmann et al. (2018) met for about 200 hours total over six months; and the empowerment and livelihood clubs in Uganda evaluated by Bandiera et al. (2020) were open five afternoons per week for two years (over 500 hours).

[5]The majority male staff was due to the very restrictions on women's mobility that the program aimed to change; the job entailed traveling across a geographic area spanning 10 treatment schools. We lack the statistical power to shed light on heterogeneous effects by facilitator gender. For both boys and girls, messages about gender equality could be more powerful coming from a man, or conversely from a woman. It is also plausible that students are most responsive to messages from same-gender instructors.

[6]This curriculum was delivered to the two study cohorts, finishing in the middle of the 2016-17 school year. Breakthrough received additional funding and resumed activities in 59 of the 150 treatment schools in 2017-18, offering an extra module to the younger of our two study cohorts. This added 0.2 years of dosage on average (59/150 * 52% of sample in younger cohort * 1 year), or 2.1 additional sessions per participant. Breakthrough also initiated the full curriculum with new cohorts, 2 to 5 years younger than our study cohorts, in these schools. We do not expect spillovers from these younger children to our study participants to have added much dosage, especially since most of the older cohort had switched to a new school for grade 11 by then. We became aware of these additional activities in 2019.

roles and stereotypes, and recognition and tolerance of discrimination. For example, one session focused on household chores. Students broke out into groups and listed whether males or females did various chores in their households. They then reconvened and discussed the answers. When the pattern emerged that women and girls did most of the chores, the facilitator asked why that was and whether it was fair. The class discussed why women cook at home, but men are cooks in restaurants, with the latter role earning more status in society. A few of the sessions aimed to impart skills such as public speaking, communication between the genders, and leadership, which could enable gender-equitable attitudes to translate into behavioral change. For instance, girls might be able to negotiate greater independence with their parents, leading to more freedom of movement in the short run and greater occupational choice in the long run. Through this curriculum, students explored gender identity and stereotypes, gained a better understanding of gender inequities and their consequences, understood their rights, and were encouraged to communicate and act on what they had learned. To map this to standard concepts used in economics, the intervention aimed to change students' preferences (i.e., their moral views on gender inequality); their factual beliefs (e.g., greater realization that restricting women's employment leaves money on the table); and their skills (e.g., how to persuade their parents to let them go to college).

II. Study design and data

A. Experimental design

We conducted a randomized evaluation of the gender attitude change program in a sample of 314 government schools across Sonipat, Panipat, Rohtak, and Jhajjar districts in the state of Haryana, India.[7] The unit of randomization was the school.

The sample size of 314 schools was chosen to be able to measure the short- and medium-run effects of the program on gender attitudes, aspirations, and behavior, as well as long-term effects on educational attainment, occupational choice, marriage, and fertility that might emerge up to ten years after the program ended. There were 607 government-run secondary schools that offered grades 6 through 10 across the four districts. We first restricted attention to the 346 schools that officially enrolled at least 40 students in grades 6 and 7 combined,

---

[7]The government prioritized these districts based on their skewed sex ratios. The 2011 child sex ratio was 1.25 in Sonipat, 1.19 in Panipat, 1.22 in Rohtak, and 1.28 in Jhajjar. These districts do not necessarily have more male-biased preferences than the rest of Haryana or north India, but their low fertility rate (they are near New Delhi) means that son preference translates into a higher rate of sex-selective abortions to ensure having at least one son in the family (Jayachandran, 2017).

and then eliminated schools with low actual enrollment based on a preliminary visit. In cases where a village had more than one government secondary school, we chose at most one of them for the sample to minimize the possibility of spillovers. Of the 314 schools in the sample, 59 enroll only girls, 40 enroll only boys, and the remaining 215 are co-ed. Official enrollment in these schools averages about 80 students per grade (DISE, 2011).

We randomly selected 150 of the sample schools to be in the treatment group; the remaining 164 serve as control schools. Figure 1 shows the four study districts and the schools assigned to the treatment and control groups. The randomization was stratified by district, co-ed status of the school, school size, and distance to the district headquarters. Table 1 reports baseline characteristics of schools by treatment status. The first panel confirms that the two samples are balanced on various school characteristics such as urban/rural and number of male and female students.

B.   Enrollment of study participants and baseline data collection

The baseline survey was conducted between August 2013 and January 2014, covering 14,809 students. The data collection was conducted by the Abdul Latif Jameel Poverty Action Lab, South Asia. All students (as well as parents, principals, field staff, etc.) were blind to treatment status when recruitment for the study and baseline data collection took place.

To select students for the sample, we visited the schools and distributed parental consent forms to all 6th and 7th graders who were present. These students would be in grades 7 and 8 in April 2014 when the program began at the start of the next school year. The parental consent rate was high; 84% of forms were returned. Perhaps surprisingly, the rate does not vary by gender or by village-level proxies for gender norms, such as the child sex ratio and female employment rate from the 2011 Census. Anecdotally, lack of consent was usually due to the student losing or forgetting the form.[8]

We randomly chose the study participants from among those whose parent gave consent, with a target of 45 students per school, stratified by gender and grade with a ratio of 3:2:2:2 for Female 6th:Male 6th:Female 7th:Male 7th. We included more girls than boys in the sample because there are more girls enrolled in government secondary schools, and

---

[8]Parents had to consent to their child participating in the study, but the program was added to the curriculum in treatment schools by the government, which did not offer parents a way to opt out. The classes were not held on a fixed day, so keeping a child home to avoid the classes would not have been straightforward. Anecdotally, parental complaints about the program to schools and the education department were negligible.

we sampled more grade 6 girls than grade 7 girls because we expected lower attrition for younger grades.[9] Students also needed to personally assent to participating in the study and be present at school on the baseline survey day. The 35-minute-long baseline survey took place on the school premises.

We mistakenly omitted one school from the baseline survey. This school was randomized into the treatment group, and it received the intervention. We collected endline data in the school and include it in the analysis, imputing baseline variables with the gender-specific sample average for the district.[10]

In addition, we surveyed one parent for a random 40% subsample of the students ($N = 6022$). We selected at random whether to interview the father or mother. We interviewed the parents at home, which added survey costs and is why we did not include all parents. We construct a gender attitude index for the parent based on nine attitude questions, which we use to understand how parental attitudes influence program impacts.

Table 1 summarizes baseline characteristics of the sample. The boys and girls were about 12 years old, on average. Religious and caste variables line up with the overall demographics for these districts, as reported in the Census (Govt. of India, 2011); the participants are predominantly Hindu. Mothers' average age was 36 years and fathers', 41 years. There is a high illiteracy rate for mothers, reflecting the low level of female schooling in the parents' generation. Consistent with India's low female labor force participation rate, only 29% of mothers are employed full-time. Baseline variables are balanced between the treatment and control groups. An F-test of joint significance fails to reject balance between the study arms. Appendix Table 1 shows summary statistics separately for girls and boys, which are also balanced between the treatment and control groups.

In the baseline survey, we included the Marlowe-Crowne module, designed by social psychologists to measure a person's propensity to give socially desirable responses (Crowne and Marlowe, 1960). The module asks the respondent whether he or she has certain almost saintly personality traits (e.g., "I am never irritated by people who ask favors of me"). Because the module is designed to use traits that people are unlikely to truly have, the interpretation when someone reports having more of these traits is that she has a stronger concern for social approval. A caveat is that some of the variation might reflect actual differ-

---

[9]Parents are more likely to send their sons than daughters to private schools. Because wealthier families use private schools, the boys in government schools are, thus, from poorer families than the girls, on average.

[10]We distributed consent forms to students present during a school visit just before the endline survey and then randomly chose sample students from among those with parental consent who assented to participate.

ences in possessing these desirable traits. The module was developed in the US but has been validated in several developing countries, including India (Mukherjee, 1967; Vu et al., 2011). We used a 13-item version of the original 33-item module (Reynolds, 1982). The questions are listed in the appendix. We combine the responses into an index, or social desirability score, which we use to investigate whether the treatment effect estimates are biased upward by experimenter demand effects. The social desirability score is balanced between the treatment and control groups, and its distribution by gender is shown in Appendix Figure 1.

C. Endline data collection

We conducted a first endline survey shortly after the program ended (three and a half months, on average), or about three years after the baseline survey. We then conducted a second endline two years later.

Data collection for the first endline survey occurred between November 2016 and April 2017. We resurveyed 13,943 of the 14,809 students surveyed at baseline, which corresponds to an attrition rate of 5.8%. The endline sample also includes an additional 44 students from the sample school that we mistakenly did not survey at baseline, yielding a total sample for the first endline of 13,987 students.

Appendix Table 2 shows that sample attrition does not differ significantly between the treatment and control groups for either boys or girls. In addition, attrition in the treatment versus control group is not differential by baseline outcomes.[11] Three fourths of students were surveyed at school at the first endline. Several students had moved to a different school, either in the same or a different village, or dropped out of school. These students were surveyed at home. If the student had moved to another village that was far from the survey districts, we conducted a truncated phone survey (0.1% of respondents).[12]

We conducted a second endline survey between January and July 2019, which was two to two and a half years after the intervention had ended and when the students were finishing or had just finished grades 11 and 12 (if they had not repeated a grade). The survey was

---

[11]Appendix Table 3 details the reasons for attrition, which include permanent or long-term migration, death or poor health, refusal to participate by the student or parent, not being available at the time of their appointment, and inability to track the respondent.

[12]Appendix Table 2 shows that treatment status is not significantly correlated with the survey location. Appendix Table 4 summarizes participants' schooling status at endline: 86% of girls and 76% of boys were enrolled in the same school as baseline; boys are more likely to have switched to a private school and also to have dropped out of school. The table also reports that over 85% of the treatment group was aware of the program activities.

conducted in students' homes, rather than at schools; the highest grade offered in most of the sample schools was grade 10, so the participants were scattered across various schools (or had dropped out). We again conducted truncated phone surveys for the small share of respondents who had moved. The attrition rate was 7.9%, and attrition is not significantly correlated with treatment status or correlated with baseline characteristics differentially by treatment status. The sample size for the second endline is 13,685 individuals.

D. Primary outcomes: Attitudes, aspirations, and behavior

We pre-specified three primary outcomes for the first endline: gender attitudes, girls' aspirations, and self-reported gender-related behavior.[13]

We combine 17 gender attitude variables into a variance-weighted index, following Anderson (2008). The specific variables and procedure for constructing the index were pre-specified (see the data appendix for more details). We measure gender attitudes mostly through direct questions about female and male roles and rights (e.g., whether women should work outside the home, the appropriate age of marriage for girls), plus a vignette about investing in a son's or daughter's education.

Gender attitudes are balanced between the treatment and control group at baseline (see Table 1).[14] They are also quite regressive. For example, about 80% of boys and 60% of girls believe that a woman's most important role is being a good homemaker. This pattern that girls are less likely than boys to endorse gender-discriminatory positions is seen for each of the attitude questions (see Appendix Table 5).

We measure girls' aspirations with a variance-weighted index combining 5 questions about educational and career aspirations. Girls have high aspirations relative to the actual rate of female college completion and employment in their communities. For example, at baseline, 72% of girls (compared to 77% of boys) expect to be employed and holding a white-collar job at age 25. This rate is much higher than the current employment rate of young women in India.

Arguably the most challenging outcome to measure was self-reported behavior. We

---

[13]The pre-analysis plan (PaP) for the first endline survey was posted to the AEA RCT Registry at the beginning of endline data collection in November 2016. It specified the primary outcomes and how they would be constructed, the secondary outcomes, heterogeneity analyses, and the procedure for choosing control variables. The PaP for the second endline was posted in December 2018 and specified similar information. The two PaPs and a short document that lists the ways in which we deviated from the PaPs are available at bit.ly/PaP4RAGA.

[14]The gender attitudes module at baseline was shorter than the endline module. The baseline attitude index aggregates nine variables.

focused on behaviors that are influenced by gender attitudes and gender norms, and that we expected adolescents to have some say over. We construct an index of 6 questions, asked of both boys and girls. Most of these questions are coded the same way for boys and girls (e.g., being supportive of sisters' career aspirations, level of interaction with the opposite gender), whereas household chores is coded so that, for girls, more gender-equal behavior is to do fewer chores whereas, for boys, it is to do more. We also included some questions applicable to only girls (e.g., mobility), which we do not include in the main index but examine in auxiliary analyses.

For the second endline, we again pre-specified gender attitudes, girls' aspirations, and behavior as primary outcomes. We construct the attitudes index identically between the first and second endlines, using the same questions and weights (which are based on the first endline data). This makes the magnitude of the effect directly comparable across the two waves. For aspirations and behavior, we updated the modules, as the relevant questions changed as the sample grew older, so we construct the variance-weighted indices independently for the two endlines.

We added two revealed-preference measures as primary outcomes in the second endline. The first captures girls' intent to pursue a college education. We set up a girls' scholarship program for college expenses and use as an outcome whether respondents filled out and mailed in the application, which we gave to them at the end of the survey visit and which required some time to fill out and submit. Application submission serves as an "intensive margin" measure of how intensely a girl aspires to attend college combined with her expectation that she will be allowed to attend college. The second revealed-preference measure focuses on both boys' and girls' willingness to publicly espouse a feminist position. We informed respondents about a petition to end the dowry system, with the names of signatories to be published in the local newspaper (through an advertisement we placed). The dowry system was not explicitly covered in the curriculum, though the topic likely came up in some discussions. Students were given a toll-free phone number to call to add their name to the petition. We use signing the petition as an outcome.

III.   Empirical specification

We hypothesized that the intervention would make participants' attitudes less discriminatory against females, raise girls' aspirations, and increase gender-equitable behavior. This section describes the estimation strategy used to test these hypotheses.

We estimate the following ordinary least squares regression, with one observation per student:

$$(1) \qquad Y_{ij} = \beta_0 + \beta_1 Treated_j + \beta_2 Y_{ij}^0 + \beta_3 \mathbf{X}_{ij} + \epsilon_{ij}$$

Y_{ij} is the outcome variable measured at endline for student i in school j. $Treated_j$ is a binary variable that equals 1 if the school was assigned to the treatment group, and 0 otherwise. Thus, β_1 represents the average effect of the intervention on the outcome. The outcomes are constructed so that a higher value represents more gender progressiveness, so the hypothesis is $\beta_1 > 0$.

We control for Y_{ij}^0, the baseline analogue of the outcome. The vector $\mathbf{X}_{ij}$ comprises other control variables, specifically grade-gender and district-gender fixed effects. When the outcome is an index, we include a missing flag for each component of the index.[15] We allow the error term, ϵ_{ij}, to be clustered at the school level, which is the level of randomization.

In addition, we estimate an alternative specification in which the regressors besides $Treated_j$ are chosen using the double LASSO procedure of Belloni et al. (2014). Appendix Table 6 lists the control variables selected for each outcome and the larger set of potential controls from which the double LASSO procedure chose them.

We also test for heterogeneous treatment effects along pre-specified dimensions: student gender and parents' attitudes at the first endline, and only student gender at the second endline. Given the parsimonious set of primary outcomes and heterogeneity analyses, we do not adjust the statistical inference for multiple hypothesis testing.

IV.   Short-run results

This section presents the results from the first endline survey, which was conducted three and a half months after the intervention ended.

A.   Effect on gender attitudes

Our first main finding is that the intervention made gender attitudes more progressive: Students in treatment schools have a 0.18 standard deviation higher attitude index than those in control schools ($p < 0.01$), as reported in Table 2, column 1.

---

[15]When a component of the index is missing for an observation, we construct the index using the remaining non-missing variables. When the baseline outcome Y_{ij}^0 is missing, we impute its value with the the district-gender mean.

One benchmark for the effect size is that endline gender attitudes are 0.50 standard deviations higher for girls than boys in the control group, as reported in Appendix Table 7. Thus, the treatment effect is 36% as large as the status quo gender gap in attitudes. Also, a one standard deviation increase in parent gender attitudes is associated with student gender attitudes being 0.05 standard deviations higher; the treatment effect is much larger than this.[16]

Another way to express the effect size is that the intervention's 'persuasion rate' was 15.5% (DellaVigna and Gentzkow, 2010). In other words, the intervention succeeded in converting 15.5% of gender-regressive views into support for gender equality. The persuasion rate is calculated by stacking all of the variables in the attitudes index, which are coded as binary values. In the treatment group, on average 29.9% of views are gender-regressive, compared to 35.4% in the control group.

When we decompose the attitude index into thematic sub-indices, we find that the program had the strongest effect on attitudes about employment and other equal rights for women and girls, followed by education attitudes, as reported in Appendix Table 8. The effect on gender-equitable fertility attitudes among girls is statistically significant but small (2 percentage points more likely to have gender-equal fertility preferences) and negligible and insignificant for boys; the intervention included very little discussion about fertility, in part because the government requested no discussion of sexual activity. The effects for each of the 17 variables that comprise the overall gender attitudes index are reported in Appendix Table 9.

The results described above use our main specification, with the basic set of controls. The effect of the intervention on the attitudes index is very similar (0.17 standard deviations; $p < 0.01$) when we instead select control variables with double LASSO, as shown in Appendix Table 10. To account for the possibility that attrition is endogenous to treatment, we also estimate Lee bounds on the treatment effects (Lee, 2009). The attrition-adjusted lower bound on the point estimate is 0.16, as shown in Appendix Table 11.

---

[16]In Dhar et al. (2019), we present an arguably better version of this parental attitudes benchmark, using indices constructed from the same set of questions for parents and students and collected at baseline for both groups. A one standard deviation increase in a parent's attitudes is then associated with a 0.11 standard deviation increase in the child's attitudes.

B.   Effects on girls' aspirations and on self-reported behavior

Turning to our second primary outcome, Table 2, column 2 shows that the program did not affect girls' aspirations. The average effect is 0.03 standard deviations and not significantly different from zero. Girls' aspirations were high to begin with. In addition, our measure does not capture how intensely the respondent held her aspirations, for example, how hard she would fight to be able to have a career. For these reasons, there might have been limited room for the intervention to raise the aspirations measure further.

The third primary outcome is self-reported behaviors influenced by gender attitudes. Behavior became more aligned with gender-progressive norms by 0.20 standard deviations ($p < 0.01$), as reported in Table 2, column 3. The magnitude and significance of this effect are robust to including additional control variables selected using double LASSO. Decomposing the behavior measure into sub-indices, we find that the intervention generated more interaction with the opposite sex for both boys and girls (see Appendix Table 12). It also increased boys' participation in household chores—a shift in the direction of a more gender-equal division—and their support for their female relatives' ambitions. In addition, among girls, the program led to greater mobility (e.g., walking to school alone) but had no impact on decision-making power.[17]

C.   Assessing bias due to experimenter demand effects

An important concern when interpreting changes in self-reported outcomes is that participating in the program might have made salient what the socially desirable responses to our survey questions were without changing actual views. The treatment group likely became more aware that many outsiders to their community regard support for gender equality as laudable. If the program caused participants to refrain from making gender-discriminatory statements in their daily lives even without changing their deeply held beliefs, that would still be a true program benefit. The concern is if the estimates reflect what participants are willing to say to the surveyors, in ways not reflective of how they act outside the study environment.

To investigate whether this type of experimenter demand effect is upward biasing the estimated program impacts, we construct a social desirability score for each respondent using the Marlowe-Crowne module administered at baseline. The score measures a person's general

---

[17]Appendix Tables 13 and 14 report the effects for the component variables that comprise the aspirations and behavior indices.

tendency to present herself in a socially desirable way in the survey. We test for heterogeneous treatment effects based on the social desirability score. The worrisome pattern would be if the treatment effects were driven by students with a high propensity to disingenuously give socially desirable answers and vanished for those with a low such tendency.

As shown in Table 3, the main effect of having an above-median social desirability score is positive and significant for all three primary outcomes, suggesting some upward shading of responses overall for the sample. We view this pattern as reassuring, a validation that the score is successful in capturing a respondent's propensity to shade responses. Importantly, there is no more of this shading up in the treatment group than in the control group; the interaction terms are small and insignificant. For example, the 0.19 main effect of *Treated* in column 1 implies that the program's effect on attitudes is large and significant when we focus on the subsample with lower susceptibility to experimenter demand effects. Appendix Table 15 shows that we find similar patterns if we use the continuous measure of the social desirability score instead of an indicator for an above-median score.

Assuming that the Marlowe-Crowne score is primarily measuring a person's propensity to give socially desirable answers (as opposed to his or her saintliness), these results help allay the concern that the estimates reflect experimenter demand effects.

D. Heterogeneous treatment effects by student gender and parental attitudes

For several reasons, the program could have had different effects on boys and girls. Girls might have felt more invested in the program's messages. Alternatively, the ideas presented could have been more eye-opening for boys, and they might have related more to the mostly male facilitators. We thus analyze the effects of the intervention separately for girls and boys. The results are reported in Table 4.

For gender attitudes, while the point estimate for the treatment effect is somewhat smaller for girls than boys, we cannot reject that the program had the same impact for both genders. The statistical comparison of the two effect sizes is based on estimating the pooled regression that maps to the two panels of Table 4 (i.e., every regressor is interacted with *Female*). The fact that girls started out more progressive than boys raises the issue of whether estimates of heterogeneity by gender are entangled with heterogeneity by initial attitudes. As shown in Appendix Table 16, the patterns of gender heterogeneity are similar

when we simultaneously allow for heterogeneity by baseline attitudes.[18]

For behavior, we find that the program had a significant positive impact for each gender, but a smaller impact for girls than boys. The p-value of this difference is < 0.01. One interpretation of this finding is that boys and girls can adopt gender-equal attitudes with relatively equal ease, but girls face more familial and societal constraints on translating their attitudes into behavior.

In light of the notable heterogeneity by gender, in subsequent tables we show the results separately by gender (and report inference based on interacted models). In addition, Appendix Tables 10, 18, and 19 show our main robustness checks (i.e., double-LASSO-selected controls, Lee bounds, social desirability bias) separately for girls and boys.

The second dimension of heterogeneity that we pre-specified was parental attitudes. In principle, the program could have had either larger or smaller effects for students whose home environment was more conservative. Table 5 reports this analysis, which uses the subsample for which we surveyed one of the parents at baseline. The index of parental attitudes is normalized to have a standard deviation of 1. When gender attitudes and girls' aspirations are the outcomes, the point estimates for the interaction coefficients are small compared to the main effects and insignificant (columns 1 and 2). Column 3 shows some weak evidence that the intervention engendered less behavior change among students from more progressive families, but the magnitude of this heterogeneity is small; it corresponds to the effect size differing by 0.04 standard deviations between the subsamples with above- and below-median parental attitudes (see Appendix Table 20 for the results using a binary measure of parental attitudes). Overall, there is limited evidence that parental support for gender equality either facilitates or hinders the average success of the intervention.[19]

E. Effects on secondary outcomes

In addition to our primary outcomes, we pre-specified a handful of secondary outcomes. One of them is perceptions of social norms. While the program only directly reached 100 to 200 adolescents per village, program participants might have started regarding pro-equality views as more mainstream and, thus, updated their belief about how common those views were in their community. Past research suggests that signals from institutions (Breakthrough

---

[18]The gender heterogeneity analysis is also robust to correcting for the gap in economic status between girls and boys in government schools, which is due to boys' higher rate of attending private schools as discussed in footnote 9. See Appendix Table 17, which controls for several wealth proxies in parallel to gender.

[19]In exploratory analyses, we find no heterogeneous effects by whether the school was co-ed, facilitator gender, siblings' sex composition, the sex ratio in the village, or the female employment rate in the village.

in this case) can be effective in changing subjective perceptions of norms (Tankard and Paluck, 2016).[20]

We examine parallel questions about (1) personally holding a positive gender attitude, (2) believing one's community has a positive gender norm in that domain, and (3) personally holding the positive attitude and believing the community will not oppose you if you act on it. We developed these questions by drawing on Bicchieri (2016). In addition to using these questions to assess how perceived social norms change, we use them to examine the extent to which participants view social norms as preventing them from acting on their progressive attitudes.

Among girls, the intervention made personal attitudes about female employment more progressive by 8 percentage points (Table 6, column 1) but did not significantly increase their perception that others in the community hold that gender-progressive view (column 2). In contrast, among boys, not only is there a treatment effect on their personal attitude, but there is also a significant increase in how progressive they view the community to be. Column 3 shows that, among both girls and boys, the treatment group is more likely to hold a progressive attitude and believe society will be supportive; many of those whose gender attitudes became more progressive do not expect the community norms to oppose their views. However, the estimates in column 3 are appreciably smaller than the estimates for personally holding that attitude reported in column 1: Some students whose own attitude changed think that restrictive cultural norms will hinder them from acting on it. Columns 4 to 6 show a broadly similar pattern regarding the social norm about women leaving home to attend college, with the intervention only changing boys' perceptions of social norms.

Another secondary outcome is school performance, which we examine to rule out the concern that the program hurt achievement in subjects like math and Hindi by taking away some of their instructional time. We find no impact on school performance, as reported in Appendix Table 21. The administrative exam-score data we use are aggregated at the school-grade level, so we cannot estimate results separately by gender. However, on average, girls outperform boys on secondary school exams in Haryana. This fact combined with girls' high aspirations to attend college meant that improvements in girls' school performance was not an expected impact of the program.

Results for the three other pre-specified secondary outcomes (described in more detail in the data appendix) are reported in Table 7. First, we find that the program improved

---

[20]Recent evidence from India and Saudi Arabia suggests that people overestimate their community's opposition to female employment (Bernhardt et al., 2018; Bursztyn et al., 2020).

girls' self-esteem (Rosenberg, 1965). Second, we find a small increase in awareness of gender discrimination among girls but not boys; status-quo awareness is already high. Third, we find no impact on two implicit association tests (IATs), each of which was administered to about 3,000 respondents.[21] One IAT measured how respondents associate girls' and boys' faces with positive and negative words. The other, originally developed by Beaman et al. (2009), associates men's and women's faces with market labor or domestic work. Because of challenges administering the IATs at baseline (13% of responses were invalidated because the completion time was too fast or slow), we did not make the IAT a key focus for the endline.[22]

V. Medium-run results

We next investigate whether the effects described above persisted. These results use data from the second endline survey, which was conducted 2 to 2.5 years after the program ended.

A. Effects on primary outcomes, including heterogeneity by gender

We continue to find a large and significant effect on gender attitudes in the second endline, as reported in column 1 of Table 8. The attitudes index is constructed identically to the first endline's index, so the 0.16 effect size medium-run effect is directly comparable to the 0.18 short-run effect size. The estimate is robust to including double-LASSO-selected controls (Appendix Table 23) and using Lee bounds (Appendix Table 11). The control group mean of 0.33 in the second endline indicates that attitudes became more progressive between the endlines, absent the intervention. Thus, the 11% fade-out in the treatment effect is not because the treatment group held less progressive attitudes at the second endline than at the first endline; rather, their attitudes improved less between the two waves than the control group's attitudes did.

Experimenter demand effects are somewhat less of a concern in the second endline because the intervention had ended two years earlier and so was less likely to be top of mind

---

[21] At baseline, we administered an IAT to a randomly selected 50% of the sample, and in the first endline we re-administered an IAT to this subsample, with half receiving each version of the IAT. Some endline IAT observations are missing due to technical problems with the laptops used.

[22] Another concern with IATs is that they measure not only personal attitudes but also awareness of shared cultural stereotypes. Thus, a program that discusses stereotypes could lead to a "worse" IAT score (Arkes and Tetlock, 2004).

for the treatment group. Nonetheless, it is important to investigate this potential confound and we do so using the same approach as earlier. We find that having a high propensity to give socially desirable answers is not associated with having larger treatment effects. That is, the coefficient on the interaction of *Treated* and having a high social desirability score is small and statistically insignificant, as shown in Table 9.

Turning to the results by gender, while there was no significant difference in attitude change between boys and girls in the short run, two years later there is. The medium-run effect size on boys' gender attitudes is 0.22, as shown in Table 10, column 2. This point estimate is slightly larger than their short-run effect size: There is no fade-out for boys. The effect for girls is 0.11 standard deviations and statistically significant ($p < 0.01$). This is two thirds of the short-run effect size for girls, though we cannot statistically rule out identical effect sizes across the two endlines. One conjecture is that acting on one's beliefs reinforces them, such that the gender gap in the persistence of attitude change is related to the smaller change in behavior among girls that we observed in the short run. In any case, we view this pattern as interesting and worthy of future research.

A treatment effect on girls' aspirations could have emerged between the two endlines if, in the control group, girls lowered their aspirations over time. This is not the case: As in the short run, we find no effect of the intervention on girls' stated aspirations in the medium run (Table 8, column 2).

For self-reported behavior, we find a sizable (0.23 standard deviations) and statistically significant treatment effect, pooled for boys and girls, just as we did in the first endline. Note that we updated the elements in the behavior index between the two endlines, so the effect size is not as directly comparable over time as the effect on attitudes is. The treatment effect on self-reported behavior continues to be larger for boys than girls in the medium run (Table 10, columns 3-4). Our interpretation of this pattern in the short run was that girls were more constrained by external factors in translating attitudes to behavior. Now another contributing factor is that attitude change is less sustained for girls.

We next examine treatment effects on the two revealed-preference outcomes that we introduced in the second endline. The first measure, for girls only, is applying for a college scholarship. The theory of change is that the program either made girls' desire to attend college more intense (higher aspirations) or enabled them to persuade their parents to support their goal (changed expectations about their future behavior, conditional on aspirations), making it more worth their while to complete the application. We find that the intervention

led to a marginally significant increase of 3.1 percentage points, or 8%, in the application rate ($p = 0.07$), as reported in Table 8, column 4). However, the p-value for this treatment effect increases to 0.13 in the alternative specification using double-LASSO-selected control variables (see Appendix Table 23).

The intervention could have affected scholarship applications either by strengthening girls' resolve to go to college or by enabling them to secure their parents' support. That is, it could have changed their preferences or relaxed a constraint. We conduct some exploratory heterogeneity analysis to further probe this. First, we find that the treatment effect on scholarship take-up is significantly higher for girls who had higher aspirations at baseline, as reported in Table 11, column 1. The total effect is close to four times as high among those with above-median initial aspirations compared to those with below-median aspirations (column 2). Second, the treatment effect on scholarship applications is significantly higher among the 80% of girls who, at baseline, said they had talked with their parents about their education goals (column 3). If the results had been concentrated among those with low aspirations, it would have been suggestive that girls' resolve to attend college increased. If it had been concentrated among those with low parental engagement, this would have pointed to girls securing their parents' support. Instead, these results do not clearly point to just one of these mechanisms operating. The intervention seems to have increased girls' intention to attend college by raising already-high aspirations and convincing already-engaged parents to support their daughter's goals, rather than by converting girls and families who started out lower on these dimensions.

Our final primary outcome is signing a public petition to end the dowry system. The intervention might have either made participants more opposed to the dowry system (attitude change) or reduced the cost to them of expressing their attitude (change in perceived social sanctions). In the control group, 15% of respondents called to add their names to the petition.[23] This rate is not significantly different in the treatment group, though the point estimate is in the direction of a small increase, as shown in Table 8, column 5. This positive point estimate is driven mostly by girls (see Table 10, columns 5 and 6). In the alternative specification using double-LASSO-selected controls, the treatment effect on signing the petition is marginally significant, with $p = 0.07$ (see Appendix Table 23). Thus,

---

[23]Ideally, for statistical power, the measure would have had a higher mean in the control group; the estimated effect size on self-reported attitudes corresponds to a large (38%) increase in the likelihood of signing. The requirement of needing to phone in may have been an impediment to signing. We chose not to have respondents sign in the presence of the surveyor, as that variant could still be susceptible to experimenter demand effects, which would defeat the purpose of adding this outcome.

we find weak evidence that the intervention influenced this outcome. One way to reconcile this weak evidence with the strong observed effect on self-reported opposition to the dowry system is that the self-reported attitude change is disingenuous. Another possibility is that students worried about repercussions in their family or community from a public statement of their view, and the intervention did not lower this perceived cost (enough). The analysis using the Marlowe-Crowne measure (Table 9, column 5) shows an interesting pattern for this outcome. The intervention significantly increased petition signing by 2 percentage points in the subsample with low social desirability concerns, as seen from the main effect of *Treated*. In addition, in the control group but not the treatment group, those with strong social desirability concerns were more likely to sign the petition; one reason for signing the petition may have been to look good to others.[24] These results are consistent with there being a set of people for whom social desirability concerns and genuine opposition to the dowry system instilled by the intervention were substitutes, with either being sufficient to sign the petition.[25]

B.   Effects on secondary outcomes

We re-examine perceived social norms as a secondary outcome in the second endline, with the results reported in Table 12. We continue to find that among boys but not girls, the treatment group views the community as more supportive of female employment, but there is no longer an effect on perceived norms about female education.

In addition, we continue to see a positive effect on girls' self-esteem. This result is reported in Table 13, along with results for the remaining secondary outcomes. The third secondary outcome is an index of girls' education outcomes. It is based on school enrollment, studying a STEM subject, taking classes to obtain extra skills (e.g., computer, English), and taking after-school tutoring for exam preparation. We see a marginally significant increase of 0.06 standard deviations in the index (Table 13, column 2).

We also added a set of questions on aspirations related to age of marriage and son preference. For both boys and girls, we find a modest increase in an index of these marriage

---

[24]In Appendix Table 22 we test whether the stated and revealed preference measures are less correlated with each other among those with a high social desirability score.

[25]In the appendix we present some further medium-run analyses. Appendix Table 24 assesses social desirability bias separately by gender. Appendix Tables 25 to 28 show the results for thematic sub-indices and the individual variables that comprise the indices. Based on the lack of robust heterogeneous results by parental attitudes at the first endline, we did not pre-specify it as a dimension of heterogeneity we would analyze for the second endline. Nonetheless, for completeness, we report these (null) results in Appendix Table 29.

and fertility aspirations, in the direction of more gender progressiveness (columns 3 and 4).

The final two secondary outcomes are related to sexual harassment. We asked girls about the harassment they experienced, and we asked boys if they engaged in sexual harassment and assault. Because of concern about underreporting by boys, we used an item count (or list) experiment for them, with one group receiving an extra item that asked them if the following statement is true: "In the past year, I have passed dirty comments about a girl; made dirty gestures in a girl's presence, or inappropriately touched or groped a girl." We find that the intervention increased girls' reported harassment (column 5), and though we do not find a significant effect on boys' reported engagement in harassment, the point estimate is positive (column 6). While we cannot disentangle whether these patterns are due to actual experiences or to changes in reporting, we speculate that the intervention made both boys and girls more aware of harassment, so they recognized and reported harassment at a higher rate. An important area for further work is to understand if these patterns, instead, reflect actual increases in harassment, which might arise from the greater interaction among boys and girls that the program encouraged.

VI. Conclusion

This paper studies an attempt to change people's views about how acceptable gender inequality is and to erode their support for societal norms that restrict women's and girls' opportunities. We examined whether an attitude-change intervention in schools in Haryana, India, could make adolescents' attitudes and, in turn, their behavior, less gender biased. The intervention had students discuss and think about gender norms and gender discrimination through class sessions held over two and a half school years.

We find that the intervention succeeded in making attitudes more supportive of gender equality and coaxing more gender-equal behavior. Moreover, the impacts persisted: When we re-surveyed participants two years after the program had ended, we continued to find strong effects.

North India has particularly strong gender discrimination, and gender norms are often highly dependent on the context. Thus, the exact programming would need to be adapted for other contexts, but this general approach of engaging adolescents, or even younger children, in school discussions could be a widely applicable way to change gender norms. While it might seem surprising that a series of class discussions changed views formed over many years, most adolescents had likely spent very little time thinking hard about gender inequality

prior to the program. Interrogating one's personal prejudices and the norms that prevail in one's community might be especially powerful in schools, without voluntary, self-selected participation and when participants are young enough that their views are still quite pliable.

That said, perhaps a more indirect approach would be even more effective than explicit discussions about the importance of gender equality; assigned readings could feature empowered women, or a history assignment could ask students to discuss women's changing role in society. This aspect of the program as well as other ones, such as participants' age, the dosage, and the relative emphasis on moral versus pragmatic arguments for equality, could be varied and assessed to optimize programs like this one.

One lesson from our results is the importance of including boys and men in programs aimed at altering gender norms. We find that attitude change translates into larger shifts in behavior for boys than girls. We speculate that the very problem the program aims to solve—that males have more power in society—means that they have more freedom to act on the gender-progressive views the program instilled in them. For this reason, as we look ahead, we expect a larger increase in employment for male participants' wives than for female participants. Of course, men also face familial constraints on their behavior and feel pressure to conform to traditional norms, so it is an open question whether such an effect on wives' employment will materialize. In future work, we hope to measure whether it indeed does, as well as how the program affects other adult outcomes such as higher education, age of marriage, and childbearing.

References

Abramsky, T., K. Devries, L. Kiss, J. Nakuti, N. Kyegombe, E. Starmann, B. Cundill, L. Francisco, D. Kaye, T. Musuya, et al. (2014). Findings from the SASA! study: A cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. *BMC Medicine 12*(1), 122.

Afridi, F., T. Dinkelman, and K. Mahajan (2018). Why are fewer married women working in rural India? A decomposition analysis over two decades. *Journal of Population Economics 31*, 783–818.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics 128*(2), 469–530.

Aloud, M., S. Al-Rashood, I. Ganguli, and B. Zafar (2020). Information and social norms: Experimental evidence on the labor market aspirations of Saudi women. NBER Working Paper No. 26693.

Alrababa'h, A., W. Marble, S. Mousa, and A. Siegel (2019). Can exposure to celebrities reduce prejudice? The effect of Mohamed Salah on Islamophobic behaviors and attitudes. Working paper No. 19-04, Immigration Policy Lab.

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

Arkes, H. and P. Tetlock (2004). Attributions of implicit prejudice, or "would Jesse Jackson 'fail' the implicit association test?". *Psychological Inquiry 15*(4), 257–278.

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

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2020). Women's empowerment in action: Evidence from a randomized control trial in Africa. *American Economic Journal: Applied Economics 12*(1), 210–259.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009). Powerful women: Does exposure reduce bias? *Quarterly Journal of Economics 124*(4), 1497–1540.

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

Bernhardt, A., E. Field, R. Pande, N. Rigol, S. Schaner, and C. Troyer-Moore (2018). Male social status and women's work. *AEA Papers and Proceedings 108*, 363–67.

Bicchieri, C. (2016). *Norms in the Wild: How to diagnose, measure and change social norms.* Oxford University Press.

Buchmann, N., E. Field, R. Glennerster, S. Nazneen, S. Pimkina, and I. Sen (2018). Power vs. money: Alternative approaches to reducing child marriage in Bangladesh, a randomized control trial. Working paper, Duke University.

Bursztyn, L., A. Gonzalez, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review 110*(10), 2997–3029.

Cantoni, D., Y. Chen, D. Yang, N. Yuchtman, and Y. Zhang (2017). Curriculum and ideology. *Journal of Political Economy 125*(2), 338–392.

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

Carrell, S., M. Hoekstra, and J. West (2019). The impact of college diversity on behavior toward minorities. *American Economic Journal: Economic Policy 11*(4), 159–82.

Crowne, D. and D. Marlowe (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology 24*(4), 349–354.

Dahl, G., A. Kotsadam, and D. Rooth (2021). Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams. *Quarterly Journal of Economics 136*(2), 987–1030.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: Empirical evidence. *Annual Review of Economics 2*(1), 643–669.

Dhar, D., T. Jain, and S. Jayachandran (2014). Evaluation of a school-based gender sensitization campaign in India. AEA RCT Registry. February 26. https://doi.org/10.1257/rct.72-11.0.

Dhar, D., T. Jain, and S. Jayachandran (2019). The intergenerational transmission of gender attitudes: Evidence from India. *Journal of Development Studies 55*(12), 2572–2592.

Dhar, D., T. Jain, and S. Jayachandran (2021). Data and Code for: Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E149882V1.

DISE (2011). District information system for education. `http://dise.in`. (accessed September 19, 2018).

Donovan, R. and S. Leivers (1993). Using paid advertising to modify racial stereotype beliefs. *Public Opinion Quarterly 57*(2), 205–218.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature 50*(4), 1051–1079.

Edmonds, E., B. Feigenberg, and J. Leight (2020). Advancing the agency of adolescent girls. Working paper, American University.

Fernandez, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *119*(4), 1249–1299.

Field, E., S. Jayachandran, and R. Pande (2010). Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review Papers and Proceedings 100*(2), 125–9.

Govt. of India (2011). Census of India. Available at `www.censusindia.gov.in/`.

Green, D., A. Wilke, and J. Cooper (2020). Countering violence against women by encouraging disclosure: A mass media experiment in rural Uganda. *Comparative Political Studies 53*(14), 2283–2320.

Grigorieff, A., C. Roth, and D. Ubfal (2020). Does information change attitudes toward immigrants? *Demography 57*, 1117–1143.

Gupta, J., K. Falb, H. Lehmann, D. Kpebo, Z. Xuan, M. Hossain, C. Zimmerman, C. Watts, and J. Annan (2013). Gender norms and economic empowerment intervention to reduce intimate partner violence against women in rural Côte d'Ivoire: A randomized controlled pilot study. *BMC International Health and Human Rights 13*(46).

Healy, A. and N. Malhotra (2013). Childhood socialization and political attitudes: Evidence from a natural experiment. *Journal of Politics 75*(4), 1023–1037.

Hopkins, D., J. Sides, and J. Citrin (2019). The muted consequences of correct information about immigration. *Journal of Politics 81*(1), 315–320.

Jayachandran, S. (2015). The root causes of gender inequality in developing countries. *Annual Review of Economics 7*, 63–88.

Jayachandran, S. (2017). Fertility decline and missing women. *American Economic Journal: Applied Economics 9*(1), 118–139.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review 69*(3), 576–595.

Jensen, R. and E. Oster (2009). The power of TV: Cable television and women's status in India. *Quarterly Journal of Economics 124*(3), 1057–1094.

Jha, P., R. Kumar, P. Vasa, N. Dhingra, D. Thiruchelvam, and R. Moineddin (2006). Low male-to-female sex ratio of children born in India: National survey of 1.1 million households. *367*, 211–218.

Kishor, S. and K. Gupta (2004). Women's empowerment in India and its states: Evidence from the NFHS. *Economic and Political Weekly*, 694–712.

Klasen, S. and J. Pieters (2015). What explains the stagnation of female labor force participation in urban India? *World Bank Economic Review 29*(3), 449–478.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. *Moral development and behavior: Theory, research, and social issues*, 31–53.

La Ferrara, E., A. Chong, and S. Duryea (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics 4*(4), 1–31.

Lee, D. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies 76*(3), 1071–1102.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review 111*(6), 1807–44.

Markus, H. and P. Nurius (1986). Possible selves. *American Psychologist 41*(9), 954–969.

McKelway, M. (2020). Women's employment in India: Intra-household and intra-personal constraints. Working paper, Massachusetts Institute of Technology.

Menon, S. (2021). Curriculum for a gender equal world. *Hindustan Times*. September 5.

Moss-Racusin, C., E. Pietri, E. Hennes, J. Dovidio, V. Brescoll, G. Roussos, and J. Handelsman (2018). Reducing STEM gender bias with VIDS (video interventions for diversity in STEM). *Journal of Experimental Psychology: Applied 24*(2), 236.

Mukherjee, B. N. (1967). A cross-validation of the Marlowe-Crowne social desirability scale on an Indian sample. *The Journal of social psychology 72*(2), 299–300.

Paluck, E. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology 96*(3), 574–87.

Porter, C. and D. Serra (2019). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics 12*(3), 226–254.

Pulerwitz, J., L. Hughes, M. Mehta, A. Kidanu, F. Verani, and S. Tewolde (2015). Changing gender norms and reducing intimate partner violence: Results from a quasi-experimental intervention study with young men in Ethiopia. *American Journal of Public Health 105*(1), 132–137.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review 109*(3), 774–809.

Reynolds, W. (1982). Development of reliable and valid short forms of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology 38*(1), 119–125.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Sen, A. (1990). More than 100 million women are missing. *The New York Review of Books 37*(20).

Tankard, M. and E. Paluck (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review 10*(1), 181–211.

Vu, A., N. Tran, K. Pham, and S. Ahmed (2011). Reliability of the marlowe-crowne social desirability scale in Ethiopia, Kenya, Mozambique, and Uganda. *BMC Medical Research Methodology 11*(1), 1–7.

Washington, E. (2008). Female socialization: How daughters affect their legislator fathers' voting on women's issues. *American Economic Review 98*(1), 311–332.

World Bank (2011). *World Development Indicators*. Washington, D.C.: World Bank.

Figure 1: Map of treatment and control schools within the study districts
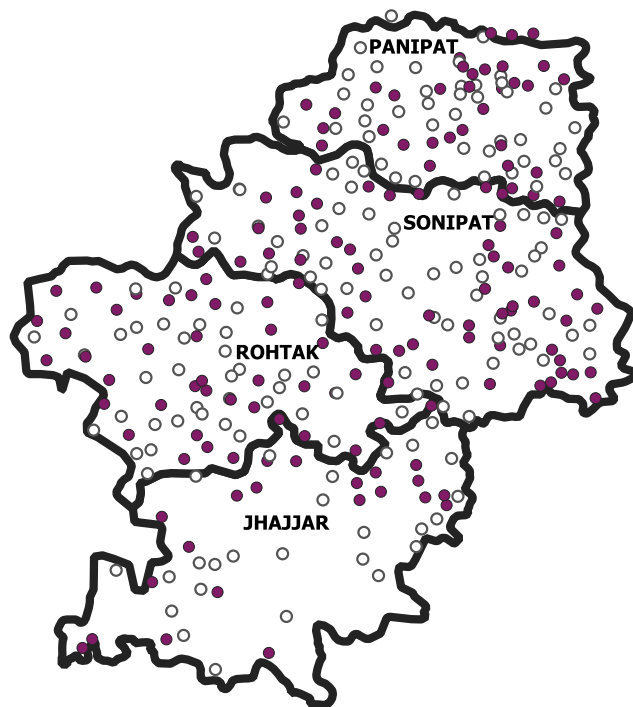
Notes: Schools in the treatment group are marked with dark dots, and schools in the control group are marked with white dots.

Table 1: Descriptive statistics: School and student characteristics at baseline

Variable	Treatment	Control	Standardized diff
Number of schools	149	164	
Urban	0.107 [0.311]	0.073 [0.261]	0.119
School is co-ed	0.698 [0.461]	0.677 [0.469]	0.045
Males in grades 6 and 7	66.427 [45.948]	65.270 [35.963]	0.028
Females in grades 6 and 7	75.125 [60.081]	74.212 [58.344]	0.015
Number of students	7,051	7,758	
Student's age	11.833 [1.261]	11.854 [1.250]	-0.017
Female	0.566 [0.496]	0.544 [0.498]	0.044
Hindu	0.945 [0.227]	0.953 [0.211]	-0.036
Enrolled in grade 6	0.526 [0.499]	0.521 [0.500]	0.011
Scheduled caste	0.268 [0.443]	0.285 [0.451]	-0.039
Mother's age	35.462 [6.351]	35.572 [6.513]	-0.017
Father's age	40.497 [6.895]	40.611 [7.131]	-0.016
Mother is illiterate	0.370 [0.483]	0.374 [0.484]	-0.009
Mother works full-time	0.292 [0.455]	0.292 [0.455]	-0.002
Dwelling has flush toilet	0.155 [0.362]	0.130 [0.337]	0.070
Gender attitudes index	0.032 [1.008]	0.000 [1.000]	0.032
Girls' aspirations index	0.039 [1.001]	0.000 [1.000]	0.039
Self-reported behavior index	-0.012 [0.994]	0.000 [1.000]	-0.012
Social desirability score	-0.025 [1.018]	0.000 [1.000]	-0.025
High social desirability score	0.374 [0.484]	0.375 [0.484]	-0.002

Notes: F-stat for joint significance of above baseline student variables is 0.924.

31

Table 2: Treatment effects on attitudes, aspirations, and behavior (Endline 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.180	0.030	0.196
	[0.020]	[0.024]	[0.021]
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	13,987	7,767	13,974

Notes: All regressions control for the baseline analogue of the outcome, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 3: Robustness check for social desirability bias (Endline 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.190	0.018	0.196
	[0.024]	[0.029]	[0.023]
High social desirability score	0.106	0.062	0.060
	[0.020]	[0.030]	[0.019]
Treated × High social desirability score	-0.024	0.032	0.001
	[0.030]	[0.043]	[0.028]
p-value: Treated + Treated × High SD = 0	0.000	0.171	0.000
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	13,987	7,767	13,974

Notes: Social desirability (SD) score is a baseline measure of the student's propensity to give socially desirable answers. High SD score refers to having an above-median score among students. All columns control for the baseline analogue of the outcome variable, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 4: Gender-specific treatment effects on attitudes, aspirations, and behavior (Endline 1)

	Gender attitudes index		Self-reported behavior index	
	Girls (1)	Boys (2)	Girls (3)	Boys (4)
Treated	0.161 [0.025]	0.204 [0.029]	0.142 [0.026]	0.260 [0.029]
Control group mean	0.237	-0.283	-0.086	0.102
Basic controls	Yes	Yes	Yes	Yes
Number of students	7802	6185	7794	6180
p-value: Girls=Boys		0.267		0.001

Notes: All regressions control for the baseline analogue of the outcome, grade and district fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 5: Heterogeneous effects by parent attitudes (Endline 1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.174	0.054	0.179
	[0.027]	[0.033]	[0.026]
Treated × Baseline parent attitudes	0.026	0.000	-0.039
	[0.025]	[0.026]	[0.021]
Control group mean	0.000	0.000	0.000
Basic controls	Yes	Yes	Yes
Number of students	5,718	3,231	5,717

Notes: All regressions control for the baseline analogue of the outcome, grade-gender and district-gender (columns 1 and 3) or grade and district (column 2) fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 6: Treatment effects on perceptions of social norms (Endline 1)

	Social norms towards work			Social norms towards education		
	Student agrees that...			*Student agrees that...*		
	women should be allowed to work	community thinks women should be allowed to work	women should be allowed to work and thinks community will not oppose them	women should be allowed to study in college even if it is far away	community thinks women should be allowed to study in college even if it is far away	women should be allowed to study in college and thinks community will not oppose them
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Girls						
Treated	0.083	0.028	0.040	0.038	0.015	0.015
	[0.011]	[0.017]	[0.016]	[0.008]	[0.018]	[0.017]
Control group mean	0.848	0.518	0.587	0.935	0.623	0.695
Number of students	3874	3661	3625	3900	3737	3717
Panel B: Boys						
Treated	0.196	0.085	0.120	0.145	0.102	0.129
	[0.020]	[0.020]	[0.020]	[0.016]	[0.019]	[0.020]
Control group mean	0.496	0.337	0.316	0.758	0.557	0.571
Number of students	2863	2691	2672	2995	2847	2833
Panel C: Girls=Boys p-value						
	0.000	0.025	0.003	0.000	0.001	0.000

Notes: All columns control for grade and district fixed effects. Each respondent was given either the set of questions on norms about work or norms about education, determined by randomization. The questions reported in columns 1 and 4, which ask about personal attitudes are not included in the gender attitudes index. Standard errors are clustered by school.

Table 7: Treatment effects on other secondary outcomes (Endline 1)

	Girls' self-esteem	Awareness of gender-based discrimination	IAT: Associates girls with positive words	IAT: Associates women with market work
	(1)	(2)	(3)	(4)
Panel A: Girls				
Treated	0.104	0.053	-0.006	-0.079
	[0.023]	[0.021]	[0.047]	[0.074]
Control group mean	0.000	0.099	0.408	0.000
Number of students	7788	7777	1676	1830
Panel B: Boys				
Treated	n/a	0.007	0.014	-0.004
	n/a	[0.020]	[0.048]	[0.063]
Control group mean	n/a	-0.118	-0.514	-0.000
Number of students	n/a	6162	1250	1368

Notes: All regressions control for grade and district fixed effects. All columns except column 2 also control for the baseline analogue of the outcome. Columns 1 and 2 also control for missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 8: Treatment effects on attitudes, aspirations, and behavior (Endline 2)

	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship	Signed petition
	(1)	(2)	(3)	(4)	(5)
Treated	0.160	-0.025	0.227	0.031	0.012
	[0.019]	[0.019]	[0.025]	[0.017]	[0.009]
Control group mean	0.333	0.000	0.000	0.408	0.150
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	13,679	7,560	13,677	7,347	13,303

Notes: All regressions control for grade-gender and district-gender fixed effects (columns 1, 3, and 5) or grade and district fixed effects (columns 2 and 4). Columns 1 to 3 also control for the baseline analogue of the outcome and missing flags for each variable used to construct the outcome index. The outcomes in columns 4 and 5 are binary variables, not indices, and were not collected for the 3% of respondents who were surveyed by phone for the second endline (because these outcomes involved giving printed material to the respondent). Standard errors are clustered by school.

Table 9: Robustness check for social desirability bias (Endline 2)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)	Applied to scholarship (4)	Signed petition (5)
Treated	0.150 [0.024]	-0.034 [0.024]	0.235 [0.028]	0.034 [0.019]	0.020 [0.010]
High social desirability score	0.070 [0.022]	0.029 [0.025]	0.059 [0.024]	0.017 [0.017]	0.014 [0.008]
Treated × High social desirability score	0.028 [0.031]	0.024 [0.034]	-0.021 [0.034]	-0.006 [0.025]	-0.021 [0.013]
p-val: Treated + Treated × High SD = 0	0.000	0.728	0.000	0.230	0.946
Control group mean	0.333	0.000	0.000	0.406	0.150
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	13,679	7,560	13,677	7,347	13,303

Notes: Social desirability (SD) score is a baseline measure of the student's propensity to give socially desirable answers. High SD score refers to having an above-median score among students. All regressions control for grade-gender and district-gender fixed effects (columns 1, 3, and 5) or grade and district fixed effects (columns 2 and 4). Columns 1 to 3 also control for the baseline analogue of the outcome and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Table 10: Gender-specific treatment effects on attitudes, aspirations, and behavior (Endline 2)

	Gender attitudes index		Self-reported behavior index		Signed petition	
	Girls (1)	Boys (2)	Girls (3)	Boys (4)	Girls (5)	Boys (6)
Treated	0.111 [0.025]	0.218 [0.028]	0.158 [0.025]	0.311 [0.040]	0.019 [0.013]	0.003 [0.010]
Control group mean	0.562	0.063	-0.067	0.079	0.189	0.104
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of students	7562	6117	7563	6114	7347	5956
p-value: Girls=Boys	0.003		0.000		0.302	

Notes: All regressions control for grade and district fixed effects, the baseline analogue of the outcome, and missing flags for each variable used to construct the outcome index. The outcome in columns 5 and 6 is a binary variable, not an index, and was not collected for the 3% of respondents who were surveyed by phone for the second endline (because these outcomes involved giving printed material to the respondent). Standard errors are clustered by school.

Table 11: Unpacking the treatment effect on scholarship applications (Endline 2)

| | Applied to scholarship | | |
	(1)	(2)	(3)
Treated	0.029	0.014	-0.023
	[0.017]	[0.019]	[0.027]
Treated × BL aspirations index	0.022		
	[0.011]		
Treated × Above-median BL aspirations		0.040	
		[0.024]	
Treated × Has discussed educ goals with parent			0.068
			[0.028]
p-value: Treated + Treated × Above-median aspir. = 0		0.020	
p-value: Treated + Treated × Has discussed goals = 0			0.016
Control group mean	0.408	0.408	0.408
Number of students	7,347	7,347	7,347

Notes: All regressions include grade and district fixed effects, the main effects for the baseline variable used in the interaction term, and flags for whether the baseline variable is missing. Standard errors are clustered by school.

Table 12: Treatment effects on perceptions of social norms (Endline 2)

	Social norms towards work			Social norms towards education		
	Student agrees that...			*Student agrees that...*		
	women should be allowed to work	community thinks women should be allowed to work	women should be allowed to work and thinks community will not oppose them	women should be allowed to study in college even if it is far away	community thinks women should be allowed to study in college even if it is far away	women should be allowed to study in college and thinks community will not oppose them
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Girls						
Treated	0.013	0.005	0.006	0.011	-0.009	-0.011
	[0.006]	[0.019]	[0.018]	[0.008]	[0.019]	[0.017]
Control group mean	0.965	0.643	0.707	0.950	0.649	0.712
Number of students	3590	3435	3418	3542	3403	3378
Panel B: Boys						
Treated	0.119	0.070	0.092	0.051	0.027	0.038
	[0.016]	[0.019]	[0.019]	[0.012]	[0.017]	[0.017]
Control group mean	0.747	0.576	0.577	0.866	0.708	0.719
Number of students	3043	2945	2935	2902	2808	2801
Panel C: Girls=Boys p-value						
	0.000	0.007	0.000	0.006	0.148	0.048

Notes: All columns control for grade and district fixed effects. Standard errors are clustered by school.

Table 13: Treatment effects on other secondary outcomes (Endline 2)

	Girls' self-esteem	Girls' education	Marriage and fertility aspirations (Girls)	Marriage and fertility aspirations (Boys)	Girls' experienced harassment	Boys' perpetrated harassment (school-grade level)
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.086	0.058	0.052	0.047	0.063	0.060
	[0.026]	[0.033]	[0.029]	[0.028]	[0.029]	[0.062]
Control group mean	0.000	0.000	0.143	0.169	0.000	-0.003
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	7,341	7,566	7,369	5,919	7,314	504

Notes: The unit of observation is the student in columns 1–5 and the school-grade in column 6. All columns control for grade and district fixed effects and, for columns 1 to 4, missing flags for each variable used to construct the outcome index. Column 1 also controls for the baseline analogue of the outcome variable. Standard errors are clustered by school. A higher value of the outcome in column 5 corresponds to higher reported harassment by girls. The outcome in column 6 is the proportion of boys who report engaging in sexual harassment, based on a list experiment.

ONLINE APPENDIX

Reshaping Adolescents' Gender Attitudes:
Evidence from a School-Based Experiment in India

Diva Dhar

University of Oxford

diva.dhar@bsg.ox.ac.uk

Tarun Jain

Indian Institute of Management Ahmedabad

tarunj@iima.ac.in

Seema Jayachandran

Northwestern University

seema@alum.mit.edu

Appendix Figure 1: Distribution of social desirability score

Panel (a): Normalized score
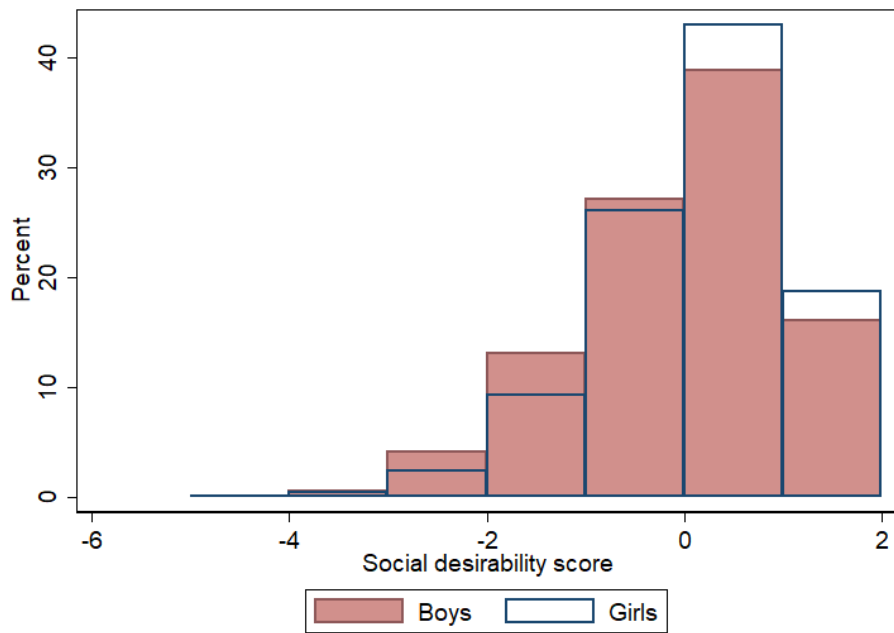
Panel (b): Raw score
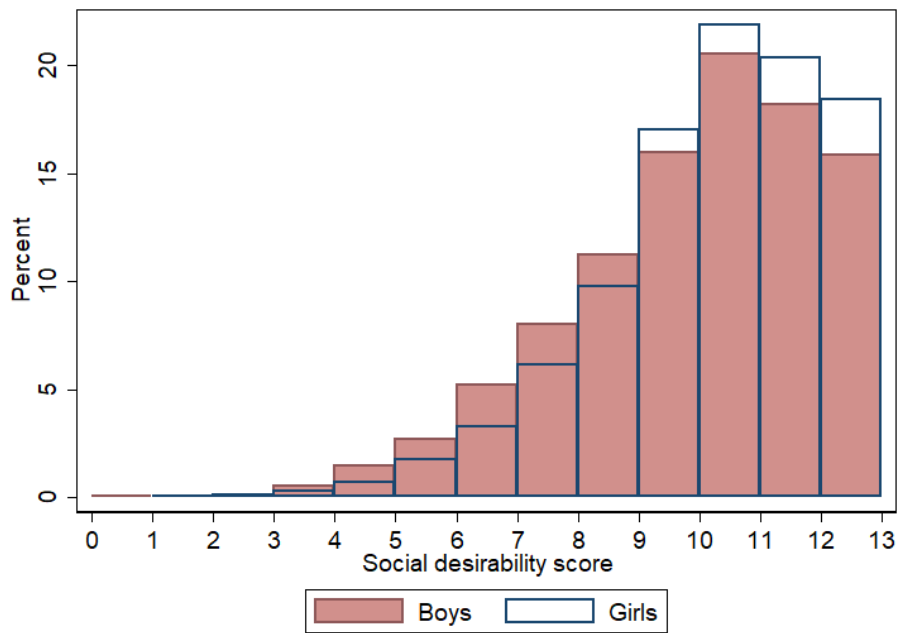
Appendix Table 1: Baseline characteristics and balance by gender

	Girls			Boys		
Variable	Treat	Control	Standardized diff	Treat	Control	Standardized diff
Number of students	3990	4222		3061	3536	
Student's age	11.712 [1.246]	11.756 [1.225]	-0.036	11.991 [1.261]	11.972 [1.271]	0.015
Hindu	0.943 [0.231]	0.955 [0.207]	-0.055	0.948 [0.222]	0.951 [0.216]	-0.014
Enrolled in grade 6	0.552 [0.497]	0.547 [0.498]	0.010	0.493 [0.500]	0.489 [0.500]	0.008
Scheduled caste	0.258 [0.437]	0.270 [0.444]	-0.027	0.281 [0.450]	0.303 [0.460]	-0.048
Mother's age	35.542 [6.535]	35.477 [6.588]	0.010	35.354 [6.094]	35.689 [6.420]	-0.053
Father's age	40.380 [6.964]	40.439 [7.245]	-0.008	40.655 [6.802]	40.821 [6.986]	-0.024
Mother is illiterate	0.365 [0.482]	0.356 [0.479]	0.019	0.376 [0.485]	0.397 [0.489]	-0.043
Mother works full-time	0.289 [0.453]	0.294 [0.456]	-0.011	0.295 [0.456]	0.290 [0.454]	0.011
Dwelling has flush toilet	0.175 [0.380]	0.151 [0.358]	0.065	0.129 [0.335]	0.106 [0.308]	0.072
Gender attitudes index	0.293 [0.902]	0.257 [0.902]	0.040	-0.308 [1.037]	-0.307 [1.025]	-0.001
Girls' aspirations index	0.039 [1.001]	0.000 [1.000]	0.039			
Self-reported behavior index	-0.604 [0.698]	-0.634 [0.680]	0.044	0.759 [0.766]	0.757 [0.767]	0.003
Social desirability score	0.051 [0.971]	0.077 [0.960]	-0.027	-0.124 [1.068]	-0.092 [1.038]	-0.030
High social desirability score	0.396 [0.489]	0.397 [0.489]	-0.002	0.345 [0.475]	0.348 [0.476]	-0.006

Appendix Table 2: Testing for differential attrition and endline survey location

Panel A: Girls

	Endline 1			Endline 2	
	Attrited	Surveyed in school	Surveyed student in-person	Attrited	Surveyed student in-person
	(1)	(2)	(3)	(4)	(5)
Treated	0.010	-0.005	-0.000	0.010	-0.006
	[0.009]	[0.018]	[0.001]	[0.011]	[0.005]
Treated × Gender attitudes index	-0.006	-0.016	0.001	-0.002	0.001
	[0.006]	[0.011]	[0.001]	[0.007]	[0.004]
Treated × Girls' aspirations index	0.003	0.018*	0.001	0.011	-0.002
	[0.005]	[0.009]	[0.001]	[0.007]	[0.004]
Treated × Self-reported behavior index	0.011	-0.025*	0.002	0.005	-0.006
	[0.007]	[0.014]	[0.001]	[0.009]	[0.005]
Gender attitudes index	0.002	0.025***	-0.000	-0.005	0.000
	[0.004]	[0.008]	[0.001]	[0.005]	[0.003]
Girls' aspirations index	-0.011***	0.005	-0.000	-0.016***	0.004*
	[0.004]	[0.006]	[0.000]	[0.004]	[0.003]
Self-reported behavior index	-0.008	0.024***	-0.001	-0.005	0.006*
	[0.005]	[0.009]	[0.001]	[0.006]	[0.003]
Control group mean	0.055	0.797	0.999	0.080	0.972
Treatment group mean	0.055	0.803	0.998	0.087	0.970
p-value: Treatment = Control	0.886	0.610	0.355	0.479	0.682
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	8,212	7,802	7,802	8,212	7,566

Panel B: Boys

	Endline 1			Endline 2	
	Attrited	Surveyed in school	Surveyed student in-person	Attrited	Surveyed student in-person
	(1)	(2)	(3)	(4)	(5)
Treated	0.004	-0.040*	0.000	0.010	-0.005
	[0.011]	[0.021]	[0.001]	[0.011]	[0.005]
Treated × Gender attitudes index	-0.005	-0.014	0.001	-0.001	0.001
	[0.006]	[0.012]	[0.000]	[0.006]	[0.004]
Treated × Self-reported behavior index	0.011	0.012	-0.000	0.006	0.002
	[0.008]	[0.015]	[0.001]	[0.009]	[0.005]
Gender attitudes index	-0.005	0.015*	-0.001	-0.003	0.001
	[0.004]	[0.008]	[0.000]	[0.004]	[0.003]
Self-reported behavior index	-0.011**	0.002	0.000	-0.005	-0.000
	[0.005]	[0.010]	[0.001]	[0.006]	[0.003]
Control group mean	0.056	0.709	0.999	0.066	0.975
Treatment group mean	0.070	0.683	0.999	0.080	0.972
p-value: Treatment = Control	0.112	0.132	0.779	0.093	0.329
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	6,597	6,185	6,185	6,597	6,119

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All columns control for grade and district fixed effects. Standard errors are clustered by school. The sample for columns 1 and 4 is the baseline sample, and the sample for columns 2, 3, and 5 are those who were successfully surveyed in the relevant endline round.

Appendix Table 3: Reasons for attrition from the sample

Panel A: Endline 1

	Girls	Boys	Total
Tracked and surveyed			
In school	6,241	4,312	10,553
At home	1,547	1,868	3,415
Over the phone	14	5	19
Tracked but could not be surveyed			
Student deceased or unwell	24	26	50
Student or parent refused assent	42	39	81
Not tracked			
Address not trackable	150	113	263
Student not found at home	15	31	46
Family in village but student moved	48	12	60
Family and student moved	113	102	215
Other	62	89	151

Panel B: Endline 2

	Girls	Boys	Total
Tracked and surveyed			
In-person	7,347	5,956	13,303
On phone	219	163	382
Parent survey*	182	116	298
Tracked but could not be surveyed			
Student deceased or unwell	43	22	65
Student or parent refused assent	189	134	323
Not tracked	276	206	484

Notes: The sample analyzed in this table are the 14,853 potential endline respondents (baseline respondents plus 44 students enrolled in the school with missing baseline data). * For Endline 2, we collected some data about students from their parents if the student was unavailable; these observations are considered to be in the "attrited" sample, as the main outcome variables are missing for these respondents.

Appendix Table 4: Descriptive statistics on school enrollment at endline

	Endline 1		Endline 2	
	Girls	Boys	Girls	Boys
Same school as at baseline	0.859	0.762	0.521	0.448
	[0.348]	[0.426]	[0.500]	[0.497]
Private school in same village/town as at baseline	0.024	0.051	0.041	0.068
	[0.152]	[0.221]	[0.199]	[0.252]
Govt school in different village/town than at baseline	0.033	0.039	0.086	0.100
	[0.179]	[0.195]	[0.280]	[0.300]
Private school in different village/town than at baseline	0.021	0.045	0.047	0.076
	[0.143]	[0.207]	[0.211]	[0.265]
Currently in formal schooling/college	0.936	0.898	0.745	0.682
	[0.244]	[0.303]	[0.436]	[0.466]
Dropped out of school and not pursing any other course	0.063	0.100	0.221	0.261
	[0.243]	[0.300]	[0.415]	[0.439]
Aware of program (treatment group only)	0.851	0.876		
	[0.356]	[0.330]		
Number of observations	7,802	6,185	7,566	6,119

Notes: Table reports variable means and standard deviations.

Appendix Table 5: Baseline attitudes and aspirations by gender

Variable	Girls	Boys
Gender attitudes index	0.274 [0.902]	-0.307 [1.030]
Disagree: A woman's most important role is being a good homemaker	0.403 [0.490]	0.201 [0.401]
Disagree: A man should have the final word about decisions in his home	0.496 [0.500]	0.334 [0.472]
Disagree: A woman should tolerate violence to keep her family together	0.665 [0.472]	0.606 [0.489]
Disagree: Wives should be less educated than their husbands	0.748 [0.434]	0.562 [0.496]
Disagree: Boys should get more opportunities/resources for education than girls	0.421 [0.494]	0.177 [0.381]
Agree: Men and women should get equal opportunities in all spheres of life	0.918 [0.274]	0.901 [0.299]
Agree: Girls should be allowed to study as far as they want	0.959 [0.199]	0.869 [0.337]
Agree: Daughters should have a similar right to inherited property as sons	0.874 [0.331]	0.823 [0.381]
Agree: It would be a good idea to elect a woman as the village Sarpanch	0.814 [0.389]	0.685 [0.465]
Self-reported behavior index	-0.620 [0.689]	0.758 [0.767]
Boys cook/clean and Girls don't	0.018 [0.133]	0.880 [0.325]
Comfortable talking to students of opp. gender	0.405 [0.491]	0.497 [0.500]
Boys take care of younger siblings and Girls don't	0.034 [0.181]	0.917 [0.275]
Aspirations index	-0.081 [1.057]	0.117 [0.931]
Student has discussed education goals with parent or adult relative	0.794 [0.405]	0.840 [0.367]
Student's highest desired level of education is above sample median	0.465 [0.499]	0.535 [0.499]
Student expects white collar job when he/she is 25 years old	0.717 [0.450]	0.772 [0.420]
Number of students	8,212	6,597

Notes: Table reports variable means and standard deviations.

Appendix Table 6: Double-LASSO-selected control variables

Extended control variable	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship (girls)	Supported petition
Student's grade at baseline					
Rural location					
Scheduled caste	X				
Scheduled tribe					
Muslim	X		X		X
Mother has completed 8th grade	BX	G			
House is pukka		G			
Dwelling has electricity					
Dwelling has flush toilet					
House has a no-flush toilet					
Family owns the house					
Household owns radio or tape recorder					
Household owns TV					
Household owns refrigerator	X				
Household gets newspaper daily	X				B
Household gets tap water					
Household owns water pump					
Self-efficacy index					
Social desirability score	BX				
Parent's baseline gender attitudes index	X		X		
Number of guest teachers in the school	G	G	G	G	G
Number of full-time teachers in the school	BGX	G	BGX	G	BGX
Fraction of female teachers	BGX		BX	G	BX
Availability of counsellor in the school	GX	G	GX	G	GX
Number of PTA meetings held in the last year	BX		BX		BX
School has a functional library	GX	G	GX	G	GX
School has functional toilets					B
School has electricity	BGX	G	BGX	G	BGX
School has access to computers	BGX	G	BGX	G	BGX
School has access to internet	GX	G	GX	G	GX
School has sports field	BX		BX	G	BGX
School has mid-day meals	BGX	G	BGX	G	BGX
School has auditorium	BGX	G	BGX	G	BGX
School has EduSat	BX		BX		BX
Bal Sabha sessions: number of times in a week	X		X		
Library sessions: number of times in a week	GX	G	GX	G	GX
School is co-educational				G	
Village-level adult female literacy rate	BGX	G	GX	G	GX
Village-level adult male literacy rate	BX		BX		BX
Village-level female labor force participation	X		X	G	GX
district-gender== 1.0000	X		X		BX
district-gender== 2.0000	X		X	G	GX
district-gender== 3.0000	X		X		X
district-gender== 4.0000			X		X
district-gender== 5.0000	X		B		X
district-gender== 6.0000	GX	G	GX	G	GX
district-gender== 7.0000	X		X		B
district-gender== 8.0000	X		X		GX
Gender attitudes index	BGX	G	GX	G	X
Girls' aspirations index	G	G	G	G	G
Self-reported behavior index	X		G		X
B-Pgender–flag					
B-Scaste-sc-flag					
B-Smuslim-flag	X		X		X
B-m-secondary-flag					
B-Shouse-pukka-y-flag					
B-Shouse-elec-flag					
B-Sflush-toilet-flag					
B-Sown-house-flag					
B-Phh-durables-1-flag					
B-Snewspaper-house-flag					B
B-Stap-water-flag					
B-q10-guest-teachr-flag	G	G	G	G	G
B-fulltime-teacher-flag	GX	G	GX	G	GX
B-pct-female-teacher-flag					
B-q13-counselor-flag	GX	G	GX	G	GX
B-q18-pta-meet-flag					B
B-q22-library-flag	GX	G	GX	G	GX
B-q22-electricity-flag	BGX	G	BGX	G	BGX
B-q22-avail-computers-flag	BGX	G	BGX	G	BGX
B-q22-avail-internet-flag	GX	G	GX	G	GX
B-q22-sports-field-flag	B		B		B
B-q22-mid-meal-flag	BGX	G	BGX	G	BGX
B-q22-auditorium-flag					
B-q22-avail-edusat-flag	BX		BX		BX
B-q21-week1-flag					
B-q21-week6-flag					

Notes: X denotes variables selected for regressions that include both genders, while G and B are those selected for the girls-only and boys-only regressions. Variable names ending in 'flag' are flags for a potential control variable having a missing value. Flags for missing components of the outcome index (when the outcome is an index) are also included in the set of potential control variables and sometimes selected but excluded from this table due to space.

Appendix Table 7: Correlates of primary outcomes in Endline 1 control group

	Gender attitudes index	
	(1)	(2)
Female	0.498***	
	[0.026]	
Baseline parent gender attitudes index		0.054***
		[0.018]
Control group mean	0.000	0.000
Basic controls	Yes	Yes
Number of students	7,326	3,003

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Sample consists of Endline 1 respondents in the control group. Column 1 controls for grade and district fixed effects, and column 2 controls for grade-gender and district-gender fixed effects. Both columns control for missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 8: Treatment effects on gender attitudes sub-indices (EL1)

	Education attitudes	Employment attitudes	Other equal rights for women	Fertility attitudes
	(1)	(2)	(3)	(4)
Panel A: Girls				
Treated	0.104***	0.125***	0.166***	0.021***
	[0.026]	[0.023]	[0.026]	[0.007]
Control group mean	0.228	0.276	0.197	0.906
Number of students	7801	7802	7802	7472
Panel B: Boys				
Treated	0.160***	0.280***	0.226***	0.009
	[0.030]	[0.030]	[0.032]	[0.010]
Control group mean	-0.273	-0.329	-0.236	0.846
Number of students	6182	6182	6182	6036

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for grade and district fixed effects and missing flags for each variable used to construct the outcome index. All columns except column 4 also include the baseline analogue of the outcome. The outcome in column 4 is not an index but a single variable that ranges from 0 to 1. Standard errors are clustered by school.

Appendix Table 9: Treatment effects on individual gender attitudes (with Bonferroni correction)

	Endline 1		Endline 2	
	Girls	Boys	Girls	Boys
Disagree: Wives should be less educated than their husbands	0.033*	0.049***	0.020	0.071***
Disagree: Boys should get more opportunities/resources for education	0.014	0.038*	0.017	0.043
If HH head, would send both children or girl for education	0.050***	0.083***	0.022	0.068***
Disagree: Woman's most important role is caring for home and children	0.100***	0.117***	0.050***	0.096***
Disagree: Men are better suited than women to work outside the house	0.081***	0.079***	0.054***	0.125***
Disagree: Marriage is more important for Pooja than her job	0.004	0.040	0.029	0.059**
Disagree: Being a teacher would be a more suitable job for Pooja	0.058**	-0.005	0.018	0.028
Agree: Women should be allowed to work outside home	0.078***	0.189***	0.013	0.119***
Agree: Daughters should have a similar right to inherited property as sons	0.022***	0.024***	0.003	0.015
Agree: It would be a good idea to elect a woman as the Sarpanch of your village	0.008	0.034***	0.007	0.024*
Disagree: A man should have the final word about decisions in his home	0.095***	0.104***	0.055***	0.110***
Disagree: A woman should tolerate violence in order to keep her family together	0.051***	0.074***	0.030***	0.047***
Disagree: Parents should maintain stricter control over daughters than sons	0.051***	0.068***	0.027	0.079***
Has gender equal views on getting higher education for better marriage prospects	-0.021	-0.011	0.018	0.019
Sister/female cousins/friends should be married after age 19	0.049***	0.083***	0.024	0.051***
Difference between boys and girls age to marry is less than control median	0.071***	0.025	0.037	0.046**
Disagree: Keep having children if no sons yet but not if no daughters	0.021**	0.009	0.015*	0.026**

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$, using Bonferroni-adjusted p-values, i.e., the raw p-value is multiplied by 17. All regressions control for grade and district fixed effects and the baseline gender attitudes index. Standard errors are clustered by school.

Appendix Table 10: Treatment effects on primary outcomes with double-LASSO controls (EL1)

	Both genders		Girls			Boys	
	Gender attitudes index (1)	Self-reported behavior index (2)	Gender attitudes index (3)	Girls' aspirations index (4)	Self-reported behavior index (5)	Gender attitudes index (6)	Self-reported behavior index (7)
Treated	0.170***	0.194***	0.155***	0.022	0.125***	0.197***	0.263***
	[0.019]	[0.021]	[0.026]	[0.026]	[0.026]	[0.030]	[0.029]
Control group mean	0.000	0.000	0.237	0.000	-0.086	-0.283	0.102
Number of students	13987	13974	7802	7767	7794	6185	6180

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The regressions control for the double-LASSO-selected variables marked in Appendix Table 6. Standard errors are clustered by school.

Appendix Table 11: Lee bounds on treatment effects for primary outcomes

	Endline 1			Endline 2				
	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship (girls)	Signed petition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.180***	0.031	0.197***	0.160***	-0.024	0.227***	0.031*	0.012
	[0.020]	[0.024]	[0.021]	[0.020]	[0.019]	[0.025]	[0.017]	[0.009]
Treated (Lower bound)	0.163***	0.030	0.183***	0.127***	-0.039**	0.193***	0.028	0.010
	[0.019]	[0.024]	[0.021]	[0.019]	[0.019]	[0.024]	[0.017]	[0.009]
Treated (Upper bound)	0.190***	0.032	0.207***	0.176***	-0.008	0.256***	0.037**	0.024***
	[0.019]	[0.024]	[0.021]	[0.019]	[0.019]	[0.024]	[0.017]	[0.008]
Observations	13,987	7,767	13,974	13,679	7,560	13,677	7,347	13,303
Observations (Lee bounds)	13,944	7,765	13,928	13,599	7,527	13,597	7,313	13,208

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. All columns control for the baseline analogue of the outcome variable and grade-gender and and district-gender fixed effects. All regressions also control for missing flags for each variable used to construct the outcome index. Standard errors are clustered by school. The Lee bound estimates are calculated by trimming observations from either the treatment or control group, whichever has a lower rate of missing data, in order to equalize the missing rate across groups.

Appendix Table 12: Treatment effects on behavior sub-indices (EL1)

	Interaction with the opposite sex	Participation in HH chores	Supporting female relatives' ambitions	Girls' decision-making	Girls' mobility
	(1)	(2)	(3)	(4)	(5)
Panel A: Girls					
Treated	0.304***	0.003	0.019	0.017	0.026***
	[0.037]	[0.022]	[0.021]	[0.028]	[0.008]
Control group mean	-0.014	-0.247	0.287	0.000	0.908
Number of students	7485	7790	7766	7791	7510
Panel B: Boys					
Treated	0.213***	0.070**	0.276***	n/a	n/a
	[0.040]	[0.035]	[0.032]	n/a	n/a
Control group mean	0.017	0.295	-0.343	n/a	n/a
Number of students	6019	6179	6161	n/a	n/a

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for grade and district fixed effects and missing flags for each variable used to construct the outcome index. All columns except column 3 also include the baseline analogue of the outcome. Standard errors are clustered by school.

Appendix Table 13: Treatment effects on components of girls' aspirations index, with Bonferroni correction (EL1)

	Girls
Expected 10th marks > control-gender median	0.014
Highest level of education you would like to complete > control-gender median	0.009
Have you discussed your education goals with your parents or adult relatives?	0.010
Child expects white collar job when he/she is 25 years old	0.006
Suppose you were to get married right after school, would you want to continue ?	0.004

Notes: Each row corresponds to a different regression of the individual variables on treatment. All regressions control for grade and district fixed effects and the baseline girls' aspirations index. Asterisks denote significance using Bonferroni correction (divides by number of variables in index – 5 – for the critical p-values.

Appendix Table 14: Treatment effects on components of self-reported behavior index, with Bonferroni correction (EL1)

	Boys	Girls
Are you comfortable talking to students of the opposite gender who are not relatives?	0.109***	0.112***
Do you sit next to students of opposite sex in the classroom?	0.080**	0.157***
How often: Cook/Clean/Wash Clothes?	0.047***	-0.002
Student has not missed school due to household responsibilities	-0.002	0.004
Disagree: Do you discourage your sister from studying in college if it is far away?	0.111***	-0.006
Disagree: Do you discourage your sister/cousin sister to work outside home?	0.110***	0.023
Student does not help with shopping for hh provisions	-0.054***	
Student does not take care of young sibling/old people	0.031**	
Student is able to talk to parents about what work she would do in the future	-0.002	
Student takes Decision: Whether or not you will continue in school past 10th grade	0.011	
Student takes Decision: If you will work after you finish your studies	0.013	
Student takes Decision: What type of work you will do after you finish your studies	0.012	
Student takes Decision: What types of chores you do at home (cooking, cleaning etc...)	0.007	
Are you allowed to go to the school alone or with friends?	0.026***	
During last week student was not absent from school	-0.007	

Notes: Each cell corresponds to a different regression of the individual variables on treatment. All regressions control for grade and district fixed effects and the baseline self-reported behavior index. Asterisks denote significance using Bonferroni correction (divides by number of variables in index for the critical p-values, i.e. 6 for common variables, 9 for girls-only variables.

Appendix Table 15: Robustness check for social desirability bias using continuous measure

	Endline 1			Endline 2				
	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship (girls)	Signed petition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.182***	0.034	0.197***	0.162***	-0.025	0.228***	0.031*	0.012
	[0.020]	[0.028]	[0.021]	[0.019]	[0.019]	[0.024]	[0.017]	[0.009]
Social desirability score	0.057***	0.029	0.038***	0.049***	0.014	0.034***	0.008	0.005
	[0.011]	[0.018]	[0.010]	[0.012]	[0.011]	[0.012]	[0.009]	[0.004]
Treated × Social desirability score	-0.009	0.018	-0.010	0.014	0.021	-0.001	0.003	-0.007
	[0.016]	[0.025]	[0.015]	[0.016]	[0.017]	[0.017]	[0.013]	[0.006]
Control group mean	0.000	0.000	0.000	0.333	0.000	0.000	0.408	0.150
Number of students	13987	7767	13974	13679	7560	13677	7347	13303

Notes: Asterisks denote significance: $*p < .10, **p < .05, ***p < .01$. Social desirability score is a baseline measure of the student's propensity to give socially desirable answers. All columns control for the baseline analogue of the outcome variable where appropriate, grade-gender and district-gender fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 16: Heterogeneity by gender, controlling for heterogeneity by BL attitudes

Panel A: Without interaction of treatment and baseline outcome

| | Endline 1 | | Endline 2 | |
	Gender attitudes index (1)	Self-reported behavior index (2)	Gender attitudes index (3)	Self-reported behavior index (4)
Treated	0.204*** [0.029]	0.260*** [0.029]	0.225*** [0.028]	0.316*** [0.042]
Treated × Female	-0.042 [0.038]	-0.118*** [0.036]	-0.121*** [0.036]	-0.156*** [0.044]
Control group mean	0.000	0.000	0.333	0.000
Basic controls	Yes	Yes	Yes	Yes
Number of students	13,987	13,974	13,208	13,207

Panel B: With interaction of treatment and baseline outcome

| | Endline 1 | | Endline 2 | |
	Gender attitudes index (1)	Self-reported behavior index (2)	Gender attitudes index (3)	Self-reported behavior index (4)
Treated	0.209*** [0.029]	0.270*** [0.032]	0.225*** [0.029]	0.299*** [0.045]
Treated × Female	-0.053 [0.040]	-0.137*** [0.042]	-0.120*** [0.039]	-0.125** [0.054]
Treated × Baseline outcome	0.018 [0.017]	-0.014 [0.019]	-0.002 [0.017]	0.022 [0.022]
Control group mean	0.000	0.000	0.333	0.000
Basic controls	Yes	Yes	Yes	Yes
Number of students	13,987	13,974	13,208	13,207

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for grade-gender and district-gender fixed effects, the baseline analogue of the outcome and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 17: Heterogeneity by gender, controlling for heterogeneity by wealth proxies

| | Endline 1 | | Endline 2 | |
	Gender attitudes index (1)	Self-reported behavior index (2)	Gender attitudes index (3)	Self-reported behavior index (4)
Treated	0.174***	0.274***	0.139***	0.298***
	[0.052]	[0.051]	[0.050]	[0.062]
Treated × Female	-0.039	-0.122***	-0.119***	-0.145***
	[0.039]	[0.036]	[0.037]	[0.042]
Treated × Father works full-time	0.018	-0.035	0.101**	0.073
	[0.045]	[0.037]	[0.043]	[0.046]
Treated × House is pukka	0.014	-0.004	-0.017	-0.002
	[0.037]	[0.032]	[0.038]	[0.038]
Treated × Dwelling has flush toilet	-0.040	0.027	-0.020	-0.001
	[0.043]	[0.041]	[0.045]	[0.051]
Treated × Household gets newspaper daily	0.061	-0.028	0.041	0.013
	[0.050]	[0.041]	[0.046]	[0.055]
Treated × Household owns some land	-0.003	0.074	0.053	-0.221***
	[0.056]	[0.047]	[0.056]	[0.057]
Father works full-time	0.036	0.012	0.011	-0.033
	[0.040]	[0.036]	[0.034]	[0.042]
House is pukka	0.025	-0.062*	-0.001	-0.072*
	[0.032]	[0.033]	[0.037]	[0.039]
Dwelling has flush toilet	0.132***	0.027	0.059	0.076
	[0.038]	[0.042]	[0.046]	[0.052]
Household gets newspaper daily	0.078	0.088*	0.123**	0.073
	[0.052]	[0.045]	[0.052]	[0.059]
Household owns some land	0.116**	-0.008	0.028	-0.141**
	[0.058]	[0.052]	[0.054]	[0.065]
Control group mean	0.000	0.000	0.333	0.000
Treatment group mean	0.204	0.216	0.521	0.223
Basic controls	Yes	Yes	Yes	Yes
Number of students	13987	13974	13679	13677

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for grade-gender and district-gender fixed effects, the baseline analogue of the outcome and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 18: Lee bounds on treatment effects, by gender

Panel A: Girls

	Endline 1			Endline 2				
	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship (girls)	Signed petition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.162***	0.031	0.142***	0.111***	-0.024	0.158***	0.031*	0.019
	[0.026]	[0.024]	[0.026]	[0.025]	[0.019]	[0.025]	[0.017]	[0.013]
Treated (Lower bound)	0.162***	0.030	0.142***	0.089***	-0.039**	0.139***	0.028	0.018
	[0.026]	[0.024]	[0.026]	[0.025]	[0.019]	[0.025]	[0.017]	[0.013]
Treated (Upper bound)	0.163***	0.032	0.143***	0.120***	-0.008	0.175***	0.037**	0.027**
	[0.026]	[0.024]	[0.026]	[0.025]	[0.018]	[0.024]	[0.017]	[0.013]
Observations	7,802	7,767	7,794	7,562	7,560	7,563	7,347	7,347
Observations (Lee bounds)	7,800	7,765	7,793	7,536	7,527	7,536	7,313	7,313

Panel B: Boys

	Endline 1		Endline 2		
	Gender attitudes index	Self-reported behavior index	Gender attitudes index	Self-reported behavior index	Signed petition
	(1)	(2)	(3)	(4)	(5)
Treated	0.204***	0.260***	0.218***	0.311***	0.003
	[0.029]	[0.029]	[0.028]	[0.040]	[0.010]
Treated (Lower bound)	0.163***	0.232***	0.173***	0.261***	0.001
	[0.029]	[0.029]	[0.027]	[0.039]	[0.010]
Treated (Upper bound)	0.231***	0.287***	0.245***	0.353***	0.020**
	[0.029]	[0.029]	[0.027]	[0.040]	[0.009]
Observations	6,185	6,180	6,117	6,114	5,956
Observations (Lee bounds)	6,138	6,131	6,065	6,063	5,897

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. All columns control for the baseline analogue of the outcome variable, grade, and district fixed effects. All regressions also control for missing flags for each variable used to construct the outcome index. Standard errors are clustered by school. The Lee bound estimates are calculated by trimming observations from either the treatment or control group, whichever has a lower rate of missing data, in order to equalize the missing rate across groups.

Appendix Table 19: Robustness check for social desirability bias, by gender (EL1)

Panel A: Girls

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.172*** [0.032]	0.018 [0.029]	0.138*** [0.028]
High social desirability score	0.074*** [0.027]	0.062** [0.030]	0.065** [0.025]
Treated × High social desirability score	-0.025 [0.040]	0.032 [0.043]	0.013 [0.035]
p-val: Treated + Treated × High SD = 0	0.000	0.171	0.000
Control group mean	0.237	0.000	-0.086
Basic controls	Yes	Yes	Yes
Number of students	7,802	7,767	7,794

Panel B: Boys

	Gender attitudes index (1)	Self-reported behavior index (2)
Treated	0.211*** [0.034]	0.259*** [0.034]
High social desirability score	0.145*** [0.034]	0.043 [0.029]
Treated × High social desirability score	-0.017 [0.051]	0.003 [0.044]
p-val: Treated + Treated × High SD = 0	0.000	0.000
Control group mean	-0.283	0.102
Basic controls	Yes	Yes
Number of students	6,185	6,180

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Social desirability (SD) score is a baseline measure of the student's propensity to give socially desirable answers. High SD score refers to having a score that is above median for the sample. All regressions control for the baseline analogue of the outcome, grade and district fixed effectsand missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 20: Heterogeneity by binary measure of baseline parent attitudes (EL1)

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)
Treated	0.199***	0.037	0.201***
	[0.038]	[0.047]	[0.035]
Treated × Above median baseline parent attitudes	-0.049	0.030	-0.043
	[0.051]	[0.056]	[0.041]
p-val: Treated + Treated × Above median attitudes = 0	0.000	0.092	0.000
Control group mean	0.237	0.000	-0.086
Basic controls	Yes	Yes	Yes
Number of students	5,718	3,231	5,717

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. All columns control for the baseline analogue of the outcome variable and grade-gender and district-gender fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 21: Treatment effects on school performance (EL1)

Panel A: SCERT school data (2014-16)

| | Proportion scoring >50 in... | | | | | |
| | Hindi | English | Math | Science | Social Science | All subjects |
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.013	-0.007	0.012	-0.020	-0.012	-0.002
	[0.018]	[0.019]	[0.018]	[0.019]	[0.018]	[0.008]
Control group mean	0.547	0.429	0.348	0.506	0.458	0.320
Control SD	0.137	0.148	0.148	0.151	0.145	0.061
Number of schools	234	230	229	228	228	237

Panel B: 10th board exam data (2017)

| | Proportion passing in... | | | | | |
| | Hindi | English | Math | Science | Social Science | All subjects |
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.001	-0.005	-0.010	-0.011	-0.021	-0.013
	[0.010]	[0.023]	[0.027]	[0.026]	[0.022]	[0.027]
Control group mean	0.924	0.667	0.751	0.735	0.733	0.552
Control SD	0.092	0.223	0.253	0.260	0.198	0.263
Number of schools	307	307	307	307	307	307

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Each observation is a school. We were able to match 237 and 307 sample schools with the SCERT and board exam datasets, respectively. The first panel uses data for both cohorts in our sample, from when each was in Grade 8. The second panel uses only the older cohort because the outcome is an exam taken in Grade 10 and the younger cohort was in Grade 9 at the time of these data. Some schools have missing observations in the SCERT dataset for certain subjects, so the sample size varies across columns within the first panel. All columns control for district fixed effects. Standard errors are heteroskedasticity-robust.

Appendix Table 22: Association between stated and revealed preferences, by social desirability score (control group only)

	Applied to scholarship		Signed petition	
	(1)	(2)	(3)	(4)
High social desirability score	-0.026 [0.039]	0.000 [0.017]	0.009 [0.017]	0.009 [0.008]
Plans to go to college	0.236*** [0.029]			
Plans college × High SDS	0.045 [0.042]			
EL2 aspirations index		0.059*** [0.012]		
EL2 aspirations index × High SDS		-0.008 [0.014]		
Against dowry			0.036*** [0.014]	
Against dowry × High SDS			0.005 [0.021]	
EL2 gender attitudes index				0.016*** [0.005]
EL2 gender attitudes index × High SDS				0.009 [0.008]
Dep var mean	0.407	0.408	0.150	0.150
Basic controls	Yes	Yes	Yes	Yes
Number of students	3,692	3,774	6,988	6,994

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. Sample includes control group only. All columns control for grade and district fixed effects. Standard errors are clustered by school.

Appendix Table 23: Treatment effects on primary outcomes with double-LASSO controls (EL2)

	Both genders			Girls					Boys		
	Gender attitudes index	Self-reported behavior index	Signed petition	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship	Signed petition	Gender attitudes index	Self-reported behavior index	Signed petition
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treated	0.153***	0.225***	0.015*	0.099***	-0.021	0.161***	0.026	0.021	0.212***	0.297***	0.007
	[0.020]	[0.025]	[0.008]	[0.027]	[0.026]	[0.025]	[0.017]	[0.013]	[0.030]	[0.042]	[0.009]
Control group mean	0.333	0.000	0.150	0.562	0.000	-0.067	0.408	0.189	0.063	0.079	0.104
Number of students	13679	13677	13303	7562	7329	7563	7347	7347	6117	6114	5956

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. The regressions control for the double-LASSO-selected variables marked in Appendix Table 6. For outcomes used in Endline 1 also (Gender attitudes index, Girls' aspirations index, and Self-reported behavior index), for consistency, we use the control variables selected by double LASSO with the Endline 1 data rather than selecting new control variables for Endline 2. Standard errors are clustered by school.

Appendix Table 24: Robustness check for social desirability bias, by gender (EL2)

Panel A: Girls

	Gender attitudes index (1)	Girls' aspirations index (2)	Self-reported behavior index (3)	Applied to scholarship (4)	Signed petition (5)
Treated	0.098***	-0.034	0.159***	0.034*	0.028*
	[0.032]	[0.024]	[0.030]	[0.019]	[0.015]
High social desirability score	0.074**	0.029	0.054**	0.017	0.018
	[0.030]	[0.025]	[0.025]	[0.017]	[0.012]
Treated × High social desirability score	0.036	0.024	-0.002	-0.006	-0.022
	[0.041]	[0.034]	[0.037]	[0.025]	[0.018]
p-val: Treated + Treated × High SD = 0	0.000	0.728	0.000	0.230	0.700
Control group mean	0.562	0.000	-0.067	0.406	0.189
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	7,562	7,560	7,563	7,347	7,347

Panel B: Boys

	Gender attitudes index (1)	Self-reported behavior index (2)	Signed petition (3)
Treated	0.211***	0.321***	0.011
	[0.033]	[0.044]	[0.011]
High social desirability score	0.066*	0.052	0.010
	[0.035]	[0.041]	[0.011]
Treated × High social desirability score	0.022	-0.028	-0.022
	[0.049]	[0.060]	[0.017]
p-val: Treated + Treated × High SD = 0	0.000	0.000	0.476
Control group mean	0.063	0.079	0.104
Basic controls	Yes	Yes	Yes
Number of students	6,117	6,114	5,956

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Social desirability (SD) score is a baseline measure of the student's propensity to give socially desirable answers. High SD score refers to having a score that is above median for the sample. All regressions control for the baseline analogue of the outcome where appropriate, and grade and district fixed effects, and missing flags for each variable used to construct the outcome index if applicable. Standard errors are clustered by school.

Appendix Table 25: Treatment effects on gender attitudes sub-indices (EL2)

	Education attitudes	Employment attitudes	Other equal rights for women	Fertility attitudes
	(1)	(2)	(3)	(4)
Panel A: Girls				
Treated	0.071***	0.121***	0.123***	0.017***
	[0.026]	[0.026]	[0.027]	[0.005]
Control group mean	0.249	0.289	0.210	0.942
Number of students	7561	7560	7560	7145
Panel B: Boys				
Treated	0.193***	0.300***	0.226***	0.026***
	[0.034]	[0.032]	[0.031]	[0.008]
Control group mean	-0.292	-0.308	-0.221	0.878
Number of students	6113	6117	6114	6006

Notes: Asterisks denote significance: $*p < .10, **p < .05, ***p < .01$. All columns control for the baseline analogue of the outcome variable (except for column 4), and grade and district fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 26: Treatment effects on self-reported behavior sub-indices (EL1)

	Interaction with the opposite sex	Participation in HH chores	Supporting female relatives' ambitions	Girls' decision-making	Girls' mobility
	(1)	(2)	(3)	(4)	(5)
Panel A: Girls					
Treated	0.282***	-0.006	0.055***	-0.006	0.137***
	[0.036]	[0.021]	[0.014]	[0.030]	[0.032]
Control group mean	0.065	-0.338	0.294	0.000	-0.000
Number of students	7331	7325	7449	7544	7544
Panel B: Boys					
Treated	0.261***	0.040	0.260***	n/a	n/a
	[0.041]	[0.029]	[0.036]	n/a	n/a
Control group mean	-0.076	0.397	-0.345	n/a	n/a
Number of students	6114	6113	6031	n/a	n/a

Notes: Asterisks denote significance: $*p < .10$, $**p < .05$, $***p < .01$. All columns control for the baseline analogue of the outcome variable (except for column 3), and grade and district fixed effects, and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

Appendix Table 27: Treatment effects on components of girls' aspirations index, with Bonferroni correction (EL2)

	Girls
Expected 12th marks > control-gender median	-0.036*
Have you ever discussed your education goals with your parents or adult relatives?	0.005
Suppose you were to get married right after school, would you want to continue ?	-0.004
Highest level of education you would like to complete > control-gender median	0.004
Reported desired occupation at age 25	-0.006
Do you plan to go to college/pursue a vocational course/pursue a professional career?	0.005
Mentioned a course she would like to pursue in higher studies	-0.013
Mentioned that she would like to pursue a STEM course in higher studies	0.010
Agree: I would like to have a job outside the home that I continue to pursue when I am married and have children	0.009
Expected or actual age of marriage	0.146
At what age would you like to have your first child?	0.077

Notes: Each row corresponds to a different regression of the individual variables on treatment. All regressions control for grade and district fixed effects and the baseline girls' aspirations index. Asterisks denote significance using Bonferroni correction (divides by number of variables in index, which is 11, for the critical p-values).

Appendix Table 28: Treatment effects on components of self-reported behavior index, with Bonferroni correction (EL2)

	Boys	Girls
Child is comfortable talking to students of the opposite sex	0.086***	0.093***
Sits next to student of opposite gender in classroom	0.062***	0.027
At least one friend of the opposite gender	0.043***	0.051***
Plays with the opposite gender	0.073***	0.084***
Talked with a child of opposite gender in past week	0.047**	0.106***
In the past one week, boy/girl did/did not cook/clean/wash dishes	0.016	-0.006
In the past month, boy/girl did/did not miss school due to household based responsibilities	0.003	0.000
Disagree: Do you discourage your sister from studying in college if it is far away?	0.046***	0.008
Disagree: Do you discourage your sister from working outside home?	0.105***	0.023***
I can talk to my parents about what work I would like to do in the future	0.002	
Child decides if studying past grade 12/vocational course	0.004	
Child decides if will work after studies	-0.006	
Child decides type of work after studies	-0.006	
Attended school every day last week	-0.016	
Are you allowed to go to the school alone or with friends?	0.015	
Have you ever gone to the market within your village to buy personal items alone?	0.061***	
Attended community events alone or with friends, no guardian	0.071***	
In the past one week, have you gone out of your house alone for any kind of purchase?	0.044*	
In past week, went to school/college alone or with friends	0.022	

Notes: Each cell corresponds to a different regression of the individual variables on treatment. All regressions control for grade and district fixed effects and the baseline self-reported behavior index. Asterisks denote significance using Bonferroni correction (divides by number of variables in index for the critical p-values, i.e. is 9 for common variables, 10 for girls-only variables).

Appendix Table 29: Heterogeneity of effects by baseline parent attitudes (EL2)

	Gender attitudes index	Girls' aspirations index	Self-reported behavior index	Applied to scholarship	Signed petition
	(1)	(2)	(3)	(4)	(5)
Treated	0.179***	-0.028	0.201***	0.026	0.008
	[0.025]	[0.028]	[0.032]	[0.023]	[0.012]
Treated × Baseline parent attitudes	-0.005	0.018	-0.020	-0.000	0.010
	[0.022]	[0.028]	[0.029]	[0.018]	[0.010]
Control group mean	0.333	0.000	0.000	0.408	0.150
Basic controls	Yes	Yes	Yes	Yes	Yes
Number of students	5429	3059	5429	2991	5309

Notes: Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. All regressions control for the baseline analogue of the outcome where applicable (columns 1-3), grade-gender and district-gender fixed effects and missing flags for each variable used to construct the outcome index. Standard errors are clustered by school.

A Data Appendix

1. Sample selection and tracking

From the universe of 607 government secondary schools in the study districts, we focused on 346 schools with enrollment of at least 40 students in grades 6 and 7 combined based on DISE (2011) data.[26] If a boys' and a girls' school shared a building or were adjacent to each other, we considered them a single school. We conducted pilot activities in 6 of the schools. We removed 26 schools from the list for the following reasons (a) if there were multiple schools on the list in a village, we randomly selected one (b) based on initial visits, actual attendance was much lower than 40 students or (c) school officials did not agree to let us conduct surveys, despite our letter of approval from the Government of Haryana. The remaining 314 schools constitute the sample.

We distributed consent forms to 30,685 students, 84% of whom returned the form signed by their parent or guardian. Most students who did not have the signed consent form when the enumerators visited said they lost it or forgot their signed form at home. Providing consent is uncorrelated with student gender, which is suggestive that it was not driven by parental gender attitudes. The school-level consent rate is uncorrelated with village-level measures of gender equality. Our sample of students for each school was randomly selected from those returning the consent form who were present on the baseline survey day in their school and assented to participate.

For the baseline parent survey, if after multiple visits and follow-up phone calls, we could not interview the selected parent, we randomly chose a replacement household. We collected data for 2,546 fathers and 3,476 mothers. The completion rate was higher for mothers than for fathers because fathers were more often unavailable during daytime hours due to work. We did not survey parents at endline.

To reduce sample attrition, we conducted two tracking surveys to verify respondents' contact information between baseline and the first endline, in January to March 2015 (98.5% tracking rate) and February to June 2016 (93.8%). We also conducted a tracking survey between the first and second endlines, in February to July 2018 (96.4%). They were conducted through a combination of in-person visits and phone calls and verified the respondent's and parents' contact information and asked about intentions to move.

2. Primary outcomes

2.1 Procedure for index construction

Most of the outcomes variables are constructed by aggregating the responses to several individual questions into an index. The index is the weighted average value of the individual variables, with weights constructed by normalizing the variables to have the same standard deviation and then recovering the weights from the inverse covariance matrix, following the procedure of Anderson (2008). The steps involved in producing the final indices are as follows:

---

[26]Because it excludes schools with low enrollment, our sample has, on average, larger villages than the universe of villages with government secondary schools. In addition to government schools, there are 731 private unaided secondary schools in the four districts, which are disproportionately in urban areas.

1. The individual variables are first converted to dummy variables. For questions that used a 5-point Likert scale, the binary variable was coded as 1 if the respondent answered "Strongly Agree" or "Agree" with a gender-progressive statement (or "Strongly Disagree" or "Disagree" with a gender-regressive statement), and 0 otherwise.

2. For the purpose of constructing the weights (but not for the final outcome variable), we impute missing values with gender-district-treatment averages. This is done to enable us to invert the covariance matrix to calculate the weights.

3. Each individual variable is normalized by subtracting the overall sample mean and dividing by the control group standard deviation.

4. Weights are generated from the inverted covariance matrix of all the normalized and imputed variables in the respective index. For some index P consisting of variables a, b, c, and d, let the weights thus produced be wt_a, wt_b, wt_c, and wt_d.

5. If an observation has missing variables (which occurs, for example, because we asked some questions to only a random subset of respondents), we construct the index using only non-missing items. We weight the non-missing variables by their respective weights and normalize by the appropriate sum of weights. For example, if a, b, and c are non-missing, the sum $S = a \times wt_a + b \times wt_b + c \times wt_c$. Let W be the sum of weights for each variable, whether missing or not. So, $W = wt_a + wt_b + wt_c + wt_d$. Let N be the sum of weights of the non-missing variables. So, $N = wt_a + wt_b + wt_c$. Then the index is calculated as $S \times (W/N)$.

6. This weighted index is then re-scaled such that the control group mean is 0 and the standard deviation is 1.

7. In our regression specifications, we control for flags for each variable in the index, indicating whether it is missing.

8. For the gender attitudes index at endline 2 (and its sub-indices), we deviate from steps 3 to 6 above. The questions in the index were identical at endline 1 and endline 2, so to make these two outcome variables identical, the component variables are adjusted by the scale factors and combined into an index using the weights that were generated with the endline 1 data. The final index is re-scaled by subtracting the endline 1 control group mean and dividing by the endline 1 control group standard deviation.

2.2 Gender attitudes index

The baseline attitude index aggregates the following 9 survey questions.

B1. A woman's most important role is being a good homemaker

B2. A man should have the final word about decisions in his home

B3. A woman should tolerate violence to keep her family together

B4. Wives should be less educated than their husbands

B5. Boys should get more opportunities/resources for education

B6. Men and women should get equal opportunities in all spheres of life

B7. Girls should be allowed to study as far as they want

B8. Daughters should have a similar right to inherited property as sons

B9. It would be a good idea to elect a woman as the village Sarpanch

The endline index aggregates 17 variables created from responses to 18 questions. Both the endlines use the questions listed here. We also divide the attitude questions into four mutually exclusive sub-indices for auxiliary analysis: gender equality in education, gender equality in employment, women's roles, and fertility preferences.

Education attitudes

E1. Wives should be less educated than their husbands

E2. Boys should be allowed to get more opportunities and resources for education than girls

E3. Education Vignette: If you were the head of the family, who would you have sent to the town for further studies?[27]

Employment attitudes

E4. A woman's most important role is to take care of her home, feeding kids and cook for her family

E5. Men are better suited than women to work outside of the house

E6. Work Vignette: Marriage is more important for Pooja than her job[28]

E7. Work Vignette: Being a teacher would be a more suitable job for Pooja

E8. Do you think women should be allowed to work outside home?

Attitudes about other equal rights for women

E9. Daughters should have a similar right to inherited property as sons

E10. It would be a good idea to elect a woman as the village Sarpanch

E11. A man should have the final word about decisions in his home

E12. A woman should tolerate violence in order to keep her family together

E13. Parents should maintain stricter control over their daughters than their sons

E14. Girls [boys] should attain higher education so that they find better husbands [wives][29]

E15. At what age would you like your sister/female cousins/friends to get married *minus* At what age would you like your brother/male cousins/friends to get married?[30]

---

[27]This question was based on a vignette about a family deciding whether to send a son or daughter to further schooling. The variable was coded as 1 if the respondent said the daughter or both children, and 0 if they answered the son.

[28]Based on a vignette about a young woman named Pooja who wants to delay marriage to pursue a job as a police officer.

[29]The variable is coded as gender progressive if the respondent gave the same responses to the question about boys and the question about girls.

[30]We code two dummies from this, the first for saying that the age for girls should be > 19 and the other for stating a gap in the appropriate age between boys and girls that was larger than the control group median response.

Fertility attitudes

E16. Suppose the first two children born to a husband and wife are both girls. Which of the following should they do? *minus* Suppose the first two children born to a husband and wife are both boys. Which of the following should they do?[31]

2.3   Aspirations index

We construct a gender aspirations index that measures educational and occupational aspirations for girls only. The questions used for the baseline aspirations index were as follows.

B1. Have you ever discussed your education goals with your parent or adult relative?

B2. What is the highest level of education you would like to complete if finances and opportunity of the school/college are available?

B3. What occupation do you expect to have when you are 25 years old?

The questions used for the aspirations index in the first endline were as follows.

E1. How many marks, according to you, will you score in the SSE 10th board examinations?

E2. Have you ever discussed your education goals with your parents or adult relatives?

E3. Suppose you were to get married right after school, would you want to continue your education after marriage?

E4. What is the highest level of education you would like to complete if finances and opportunity of the school/college are available?

E5. What occupation do you expect to have when you are 25 years old?[32]

The endline aspirations index is missing for a few observations because the respondent stopped the survey midway or refused to answer that module.

The questions used for the aspirations index in the second endline were as follows.

E2.1. How many marks, according to you, will you score in the SSE 12th board examinations?[33]

E2.2. Have you ever discussed your education goals with your parents or adult relatives?

E2.3. Suppose you were to get married right after school, would you want to continue your education after marriage?

E2.4. What is the highest level of education you would like to complete if finances and opportunity of the school/college are available? [34]

---

[31]Coded as gender regressive if the respondent said "have no more children" after having two boys but not after having two girls, and gender progressive otherwise.

[32]White collar occupations are coded as more progressive.

[33]Coded as 1 if the listed marks were greater than the gender-control group median and 0 otherwise. Question was only asked to students currently enrolled in grades 11 and 12.

[34]Coded as 1 if the level of education is greater than the gender-control group median and 0 otherwise.

E2.5. What occupation do you expect to have when you are 25 years old?[35]

E2.6. Do you plan to go to college/pursue a vocational course/professional course/join civil services or army?

E2.7. What course would you like to pursue for higher studies?[36]

E2.8. I would like to have a job outside the home that I continue to pursue when I am married and have children.

2.4 Gender behavior index

We construct a gender behavior index that measures gender equitable behavior. Questions marked with # are coded with opposite signs for boys and girls. The questions used for the baseline behavior index were as follows.

B1. Are you comfortable talking to children of the opposite gender who are not related to you inside or outside school?

B2. In the past week, did you help with cooking/cleaning/washing clothes?[#]

The endline 1 behavior index was constructed using the following questions. Questions marked with # are coded with opposite signs for boys and girls. Questions marked with ∗ were also asked in the second endline.

Interaction with the opposite sex

E1. Are you comfortable talking to children of the opposite gender who are not related to you inside and outside school?*

E2. Do you sit next to students of the opposite gender in class?*[37]

Participation in household chores

E3. In the past week, did you cook/clean/wash dishes?[#]*

E4. In the past month, have you missed school due to household based responsibilities?[#]*

Supporting female relatives' ambitions

E5. Do you discourage your sister/female cousin from working outside home?*

E6. Do you discourage your sister/female cousin from studying in college if it is far away?*

The following questions only pertain to girls and are not included in the main behavior index. These questions are included in the construction of behavior sub-indices.

Girls' decision-making

E7. I am able to talk to my parents about what work I would like to do in the future.*

---

[35]Coded as 1 if the respondent is able to report her expectations about having a job irrespective of the nature/type and 0 otherwise.

[36]Coded as 1 if the respondent is able to report any course irrespective of the nature/type and 0 otherwise.

[37]This question was not asked in single-sex schools.

The next 3 questions, E8 to E10, ask about decision-making using the following structure: "Who mostly makes decisions about the following, or if this is in the future for you, who do you expect will make this decision? Will you make the decision, make the decision jointly with parents or will parents make the decision for you?"[38]

E8. Whether or not you will continue in school past grade 10 (grade 12 in the second endline)*

E9. If you will work after you finish your studies*

E10. What type of work you will do after you finish your studies*

E11. How many days were you absent from school last week?[39]

Girls' mobility

E12. Are you allowed to travel to school alone or with friends?*

At endline 2, we excluded some behavior questions used at endline 1 (those not marked with * above) and added the following new questions to the index:

Interaction with the opposite sex

E2.1 Is friends with the opposite gender/has friends from the opposite gender.

E2.2 Plays with the opposite gender (who are not related to him/her) inside or outside of school.

E2.3 In the past one week, spoke with children (not related to him/her) of the opposite gender inside or outside of school.

The following new questions at endline 2 only pertain to girls and are not included in the behavior index. The questions are included in the 'girls' mobility' sub-index for the second endline.

Girls' mobility

E2.4. Has gone to the market within his/her village to buy personal items alone.

E2.5. Has attended community events without guardians present (either alone or with friends).

E2.6. In the past one week, has gone out of his/her house alone for any kind of purpose.

The endline behavior index is missing for a few observations because the respondent stopped the survey midway or refused to answer that module.

---

[38]Coded as 1 if the respondent alone makes the decision and 0 otherwise.

[39]Coded as 1 if the respondent was not absent to school in the previous week and 0 otherwise.

2.5 Revealed preferences measures - Endline 2 only

In the second endline, we included two revealed-preference measures. First, we offered girls an opportunity to apply for a financial scholarship to go toward college education or other post-secondary training. Second, we gave both girls and boys the opportunity to pledge support for a public petition to abolish the dowry system.

We set up a scholarship program that offered a Rs. 10,000 (150 USD) scholarship to each of 20 winners. At the end of the in-person endline survey, we informed girls about the scholarship and gave them the application form. To apply, they had to fill it out and mail it in by the stated deadline. The forms had a unique student ID, so we used the mailed-in applications to measure whether respondents applied. We randomly varied the degree of parental endorsement required on the application. Half of girls received a basic application on which they had to fill out basic information about themselves and the school and course they would like to pursue. The form also required a parental signature that stated that they understand the terms and conditions of the scholarship. In the second version of the form, there was an additional section that had to be filled by the parent or in consultation with the parent and had a weightier parental declaration that stated that they support their daughter's decision to attend college and apply for the scholarship. As pre-specified, we pool them in our analysis. (We do not see a difference in response rates between the two versions, in the control or treatment group.)

For the petition, at the end of the endline 2 survey, the enumerators informed respondents about a petition against dowry. We told respondent that names and villages of signatories would be printed in their local newspaper (and we then ran newspapers advertisements to do so). They were asked to call a toll-free number to register their support. We left a flyer with information on the petition text and the number to call. Due to a technical problem with the toll-free vendor, we lost 6 days' worth of data on potential calls from one phone carrier. Thus, we called those missed calls back to inquire if the respondents wanted to record their support.

3. Secondary outcomes

3.1 Social norms

The following questions were asked during both the endlines. Students were randomized to receive either Set 1 questions or Set 2 questions.

Set 1

E1. Do you think that women should be allowed to work outside home?

E2. Do you think that people in your village/community think that women should be allowed to work outside home?

E3. Do you think the community will oppose you since [if] you disagree with them (regarding women being allowed to work)?

Set 2

E1. Do you think that girls should be allowed to study in college even if it is far away?

E2. Do you think that people in your village/community think that girls should be allowed to study in college even if it is far away?

E3. Do you think the community will oppose you since [if] you disagree with them (regarding women being allowed to study in college)?

3.2 School performance - Endline 1 only

We examined academic outcomes to test if the intervention crowded out other academic instruction. We used overall pass rates and subject-wise test scores from two data sources:

- State Council of Educational Research and Training (SCERT): We were able to match 237 sample schools with the SCERT data. We have data for both cohorts in our sample, from when each was in Grade 8.

- Haryana Board of School Education: We were able to match 307 sample schools with the board exam dataset. We have data for only the older cohort of our sample because the outcome is 10th grade exam, and the younger cohort had not taken the exam at the time of this data collection.

3.3 Girls' self-esteem index

E1. On the whole, I am satisfied with myself

E2. I feel that I have a number of good qualities

E3. I am able to do things as well as most other people

3.4 Gender discrimination awareness index - Endline 1 only

E1. Do you know about female feticide and infanticide?

E2. Are female feticide and infanticide practiced in Haryana?

E3. According to you, what is the main reason for female feticide and infanticide?[40]

E4. In Haryana, are the number of girls less than the number of boys?

3.5 Implicit association tests - Endline 1 only

We use two gender-related implicit association tests as secondary outcomes. A random 50% of all student respondents took an IAT associating good-bad behavior characteristics to boys and girls during baseline. During endline, the same students were administered either the baseline IAT or a second IAT which asked them to associate gender stereotypical activities like factory work and ironing clothes to men and women. We use as outcomes the implicit association of girls with positive words for the first IAT, and of women with market work.

The IATs were created using Millisecond Software and administered on laptops. We aimed to collect each IAT for 25% of the sample but the usable sample size is smaller because, following guidelines for processing IAT data, we exclude observations that were completed very quickly or slowly (faster than 300 milliseconds or slower than 10 seconds on >10% of trials).

---

[40]Coded as 1 if any reason(s) given, 0 if respondent says "don't know" or doesn't answer.

3.6 Observed classroom behavior - post-Endline 1 only

After analysis of our Endline 1 data was complete and we had presented our results to some audiences, based on feedback, we decided to collect objectively measured gender-related behaviors in 2017. We developed and conducted three activities in the co-ed schools in our sample. The three activities aimed to measure (1) girls' participation in classroom discussions (2) students' views about girls' knowledge, and (3) interaction with opposite-gender peers in the classroom. In activity (1), a surveyor facilitated a class discussion about "What changes do you want to see in your society?" Another surveyor took note of how many girls and boys made comments in the discussion. In activity (2), students were told about an inter-school competition based on a general knowledge quiz. The winning classroom in each district would get school bags for every student in the class. Students were asked to vote for three students in their class to represent them. The outcome is how many girls are elected for the quiz competition. For activity (3), students were asked to form groups of five for a poster-making activity about "Swachh Bharat Abhiyan" (India's Cleanliness Drive). The surveyor recorded how many of the groups were mixed-gender.

There were some major limitations of this exercise. First, our pilot activities were too limited to reveal to us that there is no gender gap in class participation in the status quo (i.e., in the control group), and students do not perceive girls' knowledge as lower than boys', making the first two activities ill-suited for testing for changes in gender roles and stereotypes. Second, we have low power to detect changes in the outcomes, partly because we only received permission from principals to conduct the exercises in 197 schools. Also, for our third outcome (co-ed poster-making teams), only 5% of self-formed groups in the control group were mixed-gender, so we only have power to detect a very large proportional increase in this outcome. With those caveats, we find no significant effect on these outcomes. Results available upon request.

3.7 Girls' education index - Endline 2 only

E2.1. Which school are you enrolled in?[41]

E2.2. What stream are you currently following?[42]

E2.3. In the past one year, have you enrolled for an English speaking, computer training, or vocational class?

E2.4. Do you take after-school/college tuitions?

3.8 Marriage and fertility aspirations - Endline 2 only

E2.1. At what age do you want to marry?[43]

E2.2. At what age do you want to have your first child?[44]

---

[41] Coded as 1 if respondent is in any school or college, including open school, and 0 if dropped out or in vocational training.

[42] Coded as 1 if pursuing a science, commerce with math, or arts with math stream.

[43] Coded as 1 if the age is greater than the gender-control group median and 0 otherwise.

[44] Coded as 1 if the age is greater than the gender-control group median and 0 otherwise.

E2.3. How many children do you want to have? How many of these children would you like to be boys, how many would you like to be girls?[45]

E2.4. Suppose your spouse and you are going to have N children, how many of them would you want to be boys?[46]

E2.5. If instead of X boys and $N - X$ girls, you could either have $X - 1$ boys and $N - X + 1$ girls OR $X + 1$ boys and $N - X - 1$ girls, which would you prefer?[47]

3.9   Girls' experience of sexual harassment/assault - Endline 2 only

The index is coded so that a higher value corresponds to more instances of harassment.

E2.1. In the past one year, have you ever been slapped, hit, or otherwise physically hurt by a boy in a way you did not want?

The following questions are coded as 1 if the incidence ever occurred, and 0 otherwise.

E2.2. How frequently have you been teased, whistled at, or called names by boys in school in a way you did not want?

E2.3. ...teased, whistled at, or called names by boys outside of school in your village/town in a way you did not want?

E2.4. ...touched or groped by boys in school in a way you did not want?

E2.5. ...touched or groped by boys in your village/town in a way you did not want?

3.10   Boys' engagement in sexual harassment/assault - Endline 2 only

We asked the boys in our sample about sexual harassment/assault, using list randomization. Half the boys in the sample, stratified by treatment, were given a list of 5 questions including the sensitive question. The other half were given a list of the 4 non-sensitive questions. They were asked how many of the statements were true without having to list which statements were true. We calculate the school-grade level differences between the mean true statements in the two sets as a measure of the proportion of boys who engage in harassment/assault. This outcome is analyzed at the school-grade level. Single-sex girls schools are thus excluded when we analyze this outcome.

The statements given to the respondents are as follows, with the sensitive item italicized:

E2.1. In the past year, I have made new friends.

---

[45]Coded as 1 if the number of girls desired is greater than or equal to the number of boys and 0 otherwise.

[46]The N was randomly generated integer between 1 and 5, inclusive. Question coded as 1 if the number of girls desired is greater than or equal to the number of boys and 0 otherwise.

[47]X is the number of boys that the respondents said they desire to have out of the randomly generated N number of children. The response options of the questions are as follows: (a) Prefer $X - 1$ boys, and $N - X + 1$ girl or (b) Prefer $X + 1$ boys, and $N - X - 1$ girls. The question was coded as gender progressive if the respondent chooses response option 1 and 0 otherwise.

E2.2. *In the past year, I have passed dirty comments about a girl; made dirty gestures in a girl's presence or inappropriately touched or groped a girl.*

E2.3. In the past year, I have gone on a vacation with my parents (to a relative's place etc.)

E2.4. In the past year, I have scolded my friend/cousin.

E2.5. In the past year, I have watched a program (sports, cultural etc.) on television.

4. Social desirability score

We use a 13-question short form of the Crowne and Marlowe (1960) module developed by Reynolds (1982). The following questions were asked at baseline with two answer choices: agree or disagree. The social desirability score sums how many of the responses are the socially desirable one. A low score means a lower tendency to give answers that have social desirability bias.

B1. It is sometimes hard for me to go on with my work if I am not encouraged

B2. I sometimes feel resentful when I don't get my way

B3. On a few occasions, I have given up doing something because I thought too little of my ability

B4. There have been times when I felt like rebelling against people in authority even though I knew they were right

B5. No matter who I'm talking to, I'm always a good listener

B6. There have been occasions when I took advantage of someone

B7. I'm always willing to admit it when I make a mistake

B8. I sometimes try to get even rather than forgive and forget

B9. I am always courteous, even to people who are disagreeable

B10. I have never been irked when people expressed ideas very different from my own

B11. There have times when I was quite jealous of the good fortune of others

B12. I am sometimes irritated by people who ask favors of me

B13. I have deliberately said something that hurt someone's feelings

5. Parent's gender attitude index

To understand how parental attitudes influence program impacts, one parent of a random 40% subsample of the surveyed students participated in a survey during baseline. The following questions were used to construct our parent's gender attitudes index at baseline.

B1. A woman's most important role is being a good homemaker

B2. A man should have the final word about decisions in his home

B3. A woman should tolerate violence to keep her family together

B4. Wives should be less educated than their husbands

B5. Boys should get more opportunities/resources for education

B6. Men and women should get equal opportunities in all spheres of life

B7. Girls should be allowed to study as far as they want

B8. Daughters should have a similar right to inherited property as sons

B9. It would be a good idea to elect a woman as the village Sarpanch

The heterogeneity analysis with the parent index is restricted to students whose parents were surveyed. We also use the parent attitude index as a possible control variable in our double LASSO procedure. We impute missing values at the mean value for those students whose parent was not surveyed.

B Examples of intervention topics and activities

Below is a list of some of the activities and discussion topics that were part of the curriculum developed and implemented by Breakthrough.

- Students play a mixed gender team building exercise of "10 pass". Each mixed team tried to pass the ball 10 times between members without being caught by other team, to help establish trust and build confidence and comfort in working and interacting with the opposite gender.

- Students play "Antakshari" (an Indian singing game in which players sing snippets of popular songs) and discuss how song lyrics represent or characterize men and women differently, and analyze the influence of gendered roles on film songs.

- Students make posters in the classroom on how to achieve and create gender equality. The posters are displayed on the walls, and students walk through to discuss the ideas and content of the posters, and to vote on the best poster.

- Students read a book of stories, folk tales or fairy tales and identify gendered roles and stereotypes and gender-based discrimination within the book.

- Students read a comic story about a young girl who becomes a policewoman and addresses various stereotypes. Students reflect and discuss her story within groups, and share stories of addressing gender-based discrimination.

- Students divide themselves into two teams, and perform role play and act out situations related to gender-based discrimination (for example, a girl being subjected to speaking out against sexual harassment), which helps recognize gender stereotypes and roles and how to break them.

- Students divide into two teams and debate for and against important gender equality topics (e.g. Is higher education equally important for girls or boys? Why?) to recognize how norms and perceptions influence the gender roles they play in family and schools and the need to question where these values come from.

- Students learn to identify gender-based discrimination by playing the "Yes or No" game, where groups run to a "No" or "Yes" corner of the room in response to a question posed on gender-based discrimination such as "Should girls and women be allowed to wear what they want?" followed by a discussion on the rationale of their responses. Students are also asked to analyze the short-term and long-term impact of gender-based discrimination on boys and girls.

- Students role play exercises to practice being assertive, saying 'yes' or 'no' clearly and negotiating with peers to resist bullying and sexual harassment, and to intervene and stop peers from bullying or harassing others.

- Students build an action plan with suggestions and strategies for creating a safe environment for boys and girls in school, which is free of sexual harassment and gender-based violence. This is accompanied with an interactive exercise led by students on writing a petition to the school principal to create a safe environment for all genders, after which they present the student-led action plan to the school administration. Students then co-create a press release on action taken by the school to build a safe environment for all genders.

- Students conduct puppet shows to share positive stories of female role models and icons who have fought against gender-based discrimination and broken through glass ceilings (e.g. Kalpana Chawla, an astronaut).