

# 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.<sup>26</sup> 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:

---

<sup>26</sup>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?<sup>27</sup>

### **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<sup>28</sup>
- 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]<sup>29</sup>
- 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?<sup>30</sup>

---

<sup>27</sup>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.

<sup>28</sup>Based on a vignette about a young woman named Pooja who wants to delay marriage to pursue a job as a police officer.

<sup>29</sup>The variable is coded as gender progressive if the respondent gave the same responses to the question about boys and the question about girls.

<sup>30</sup>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?<sup>31</sup>

### 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?<sup>32</sup>

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?<sup>33</sup>
- 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? <sup>34</sup>

---

<sup>31</sup>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.

<sup>32</sup>White collar occupations are coded as more progressive.

<sup>33</sup>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.

<sup>34</sup>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?<sup>35</sup>
- 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?<sup>36</sup>
- 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?#<sup>37</sup>

### 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.\*

---

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

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

<sup>37</sup>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?”<sup>38</sup>

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?<sup>39</sup>

### **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.

---

<sup>38</sup>Coded as 1 if the respondent alone makes the decision and 0 otherwise.

<sup>39</sup>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?<sup>40</sup>
- 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).

---

<sup>40</sup>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?<sup>41</sup>

E2.2. What stream are you currently following?<sup>42</sup>

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?<sup>43</sup>

E2.2. At what age do you want to have your first child?<sup>44</sup>

---

<sup>41</sup>Coded as 1 if respondent is in any school or college, including open school, and 0 if dropped out or in vocational training.

<sup>42</sup>Coded as 1 if pursuing a science, commerce with math, or arts with math stream.

<sup>43</sup>Coded as 1 if the age is greater than the gender-control group median and 0 otherwise.

<sup>44</sup>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?<sup>45</sup>
- E2.4. Suppose your spouse and you are going to have  $N$  children, how many of them would you want to be boys?<sup>46</sup>
- 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?<sup>47</sup>

### 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.

---

<sup>45</sup>Coded as 1 if the number of girls desired is greater than or equal to the number of boys and 0 otherwise.

<sup>46</sup>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.

<sup>47</sup> $X$  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