

Can a Girl Influence Her Own Schooling in Early Adolescence? Evidence from Rajasthan

Eric Edmonds
Dartmouth College and NBER

Ben Feigenberg
University of Illinois, Chicago

Jessica Leight*
IFPRI and American University

December 20, 2019

Abstract

Using a clustered randomized control trial, this study considers the impact of a school-based life skills intervention aimed at 2,459 girls in Rajasthan, India starting in grade six. We find a 25 percent decline in school dropout that persists from seventh grade through the transition to high school. The intervention appears to improve the girls' reported level of agency and sense of social connectedness without measurable changes in academic performance. These findings highlight that even in a setting of gender disadvantage, girls can play a role in intrahousehold decision-making around school enrollment and that social ties are an important part of their decision-making.

JEL: I25, J16, O15

Keywords: Human Capital, Non-Cognitive Skills, Gender, India

*Funding is provided by the United States Department of Labor under cooperative agreement number IL-26700-14-75-K-25. This material does not necessarily reflect the views or policies of the United States Department of Labor, nor does mention of trade names, commercial products, or organisations imply endorsement by the United States Government. 100 percent of the total costs of the project or program is financed with Federal funds, for a total of \$1,304,957 dollars. This study was registered in the AEA RCT Registry ID AEARCTR-0001046. Link to registration: <https://doi.org/10.1257/rct.1046-3.0>. We are grateful to Mohar Dey, Rakesh Pandey, and Amanda Sload for their research assistance and appreciate their input into this project. Supporting materials for this project including referenced appendices are available at: <https://sites.dartmouth.edu/eedmonds/gep/>

1 Introduction

Schooling decisions are among the most important in a person’s life, and in the majority of low income countries, those decisions are made almost entirely while a person is a child, dependent on her caregivers. Models of schooling decisions in childhood typically assign decision-making entirely to parents (Becker, 1962; MacLeod and Urquiola, 2018), and the associated literature on understanding the enduring gender gap in education in low-income countries similarly focuses on factors that are outside of a child’s control: for example, son preference, labor market opportunities, and marriage norms. Yet the transition between childhood (typically modeled without agency) and adulthood (modeled with complete agency) is not discrete: children may have some agency in schooling decisions, and their preferences around schooling may accordingly be relevant for observed outcomes. Bursztyrn and Coffman (2012) document that parents in Brazil face challenges in monitoring and incentivizing school attendance among adolescents, and Berry (2015) highlights the moral hazard problem parents face with respect to child test-taking performance.

This study is motivated by the question of whether it is possible to shift non-cognitive skills among girls in a setting of gender disadvantage and, through the development of these skills (and in the absence of any relaxation of other external constraints), increase schooling. More specifically, we utilize a randomized controlled trial to evaluate a life skills curriculum and mentoring program in a sample of 2,459 adolescent girls in 119 schools in the Ajmer district of Rajasthan, India. This is a setting in which it is reasonable to hypothesize that both preferences and limited agency may be meaningful constraints for girls’ schooling. Baseline data collected for this evaluation documents that less than 25 percent of adult women work outside the home; over 90 percent of our respondents state that a wife should always obey her husband, and, at an average age of eleven, 17 percent of our subjects are married.

The intervention analyzed in this evaluation is the Girls’ Education Program (GEP), and its core elements include biweekly life skills classes conducted in school as well as group mentoring sessions for girls.¹ The program is delivered by social mobilizers, women from the area who have completed secondary school and who are managed, trained and deployed by our partner non-governmental organization, Room to Read (RtR). The stated objective of GEP is to enhance girls’ life skills and increase secondary school completion; the intervention targets girls beginning in grade six, and was newly rolled out to 60 randomly selected treatment schools as of the school year beginning in the

¹The program’s definition of mentoring is not what we expect most readers to have in mind. Mentoring sessions are small group discussions around topics covered in the life skills classes, principally led by students.

summer of 2016.

We examine the impact of GEP using an intent-to-treat empirical strategy. 96 percent of girls offered treatment attended at least one life skills class, and 85 percent of them remained engaged through grade seven. The majority of our analyses are based on household surveys collected after grade seven, when girls in treatment schools have been exposed to two full years of the program.

In our primary empirical results, we find that random assignment to GEP reduces school dropout. In the endline survey conducted following seventh grade, we document a 25 percent reduction in dropout (equivalent to a decline of three percentage points) and a parallel increase in grade progression. We also utilize administrative records from schools that extend through the initiation of grade nine and find that the reduction in school dropout continues through the progression into high school (a frequent point of dropout). While a large literature documents improvements in girls' education from a variety of material transfers such as uniforms (Duflo et al., 2015), bicycles (Muralidharan and Prakash, 2017), and cash (Dhaliwal et al., 2013; Fizbein et al., 2009; Baird et al., 2013), our dropout finding contributes to a smaller evidence base around whether non-material interventions seeking to target underlying attitudinal barriers can generate shifts in schooling, and contrasts with the evidence from the U.S. that generally does not find effects of adolescent targeted life-skills programs on dropout (Cunha et al., 2006; Levitt et al., 2016; Lavecchia et al., 2016).

We also find evidence of a significant enhancement in reported life skills after seventh grade: girls perceive improvements in social and emotional support and empowerment, articulate more positive gender norms, and exhibit evidence of increases in future planning. We present suggestive evidence that these patterns do not reflect only social desirability bias, girls simply repeating the curriculum content back to enumerators. While several studies document labor market returns to many of the attributes GEP seeks to develop (e.g. Heckman et al., 2006; Deming, 2017), the majority of the evidence around developing these skills in adolescence either combines the effects of life skills and other interventions (e.g. Buchmann et al., 2017; Bandiera et al., 2019b; Bandiera et al., 2019a), focuses on imparting one type of information (Nguyen, 2008; Jensen, 2010), or evaluates the impact of developing one class of life skills (Dhar et al., 2018; Ashraf et al., 2018). Our findings highlight the returns to developing girls' non-cognitive skills across a broad set of domains simultaneously and to using a classroom-based delivery method that takes advantage of existing infrastructure, and potentially leverages girls' social networks already developed within the school setting. It may be that the in-school setting that leverages existing social relationships is an important driver of our findings, in contrast to Delavallade et al. (2017) who target out of school girls with "peer group

learning” to little effect.

Beyond the effects on dropout and non-cognitive skills, however, we observe no significant effects on ancillary outcomes. There is no evidence of shifts in the probability of marriage, girls’ engagement in labor inside or outside the home, their time allocated to education, or their cognitive skills as measured by ASER cognitive tests administered at home.

The observed decline in dropout could be consistent with girls’ increased ability to advocate for themselves and their schooling goals vis-a-vis their parents; and/or an increased desire by girls to remain in school due to the enhanced social support they perceive in school and enhanced aspirations. Ashraf et al. (2018) evaluate an intervention in Zambia that specifically teaches negotiation skills to adolescent girls and finds that the intervention generated improved educational enrollment among girls. Girls’ improved ability to exercise their preferences could also help explain our dropout findings, but it is not obvious that girls are more inclined to want to go to school than parents desire (Bursztyn and Coffman, 2012) and we do not observe changes in many other activities where parents and children are likely to have divergent preferences (marriage, time allocation, etc.).

Rather, the evidence is more consistent with a process in which girls engaged in life skills classes form closer relationships with their female classmates; these stronger friendships, in turn, play an important role in explaining why GEP reduces dropout. In qualitative interviews, girls consistently emphasized that GEP led them to expand the breadth and depth of their social engagements and to increase the time spent doing schoolwork together rather than alone. These changes encouraged girls to continue attending school. Paralleling this qualitative finding, we observe increases in social engagement in response to treatment assignment in our data.

While many studies emphasize the importance of returns to education for schooling decisions, we do not observe significant changes in girls’ aspirations around further educational attainment or employment as we would expect if the treatment was changing girls’ enrollment rates by affecting perceptions of future returns to schooling. Treated girls also do not appear to be motivated to do well in school (they do not study more conditional on staying in school), and they do not perform better in testing. By identifying changes in girls’ sense of social support and social connectedness as potentially central drivers of estimated effects on school enrollment, our findings complement work such as Bursztyn et al. (2017) in Brazil, highlighting that social connectedness may be an important driver of schooling decisions, and Delavallade et al. (2016), who document that social exclusion can reduce educational aspirations. At the same time, while our study is part of the nascent literature that aims to understand how attitudinal changes translate into behavioral ones,

our findings are consistent with Dhar et al. (2018), which highlights a limited sphere of influence for changes in attitudes around girls in a setting of gender disadvantage.

The paper proceeds as follows. Section 2 describes the setting and Section 3 describes the evaluation design and empirical strategy. Section 4 presents the empirical findings and Section 5 presents a discussion of the central findings. Section 6 concludes.

2 Background

2.1 Setting

This evaluation was conducted in Ajmer, Rajasthan in northwest India. Rajasthan has been a focus of RtR programming in part because of persistent gender enrollment gaps. The programmatic rollout analyzed in this evaluation represented a substantial expansion for RtR in the state; RtR was active in only four schools in Ajmer prior to this evaluation, entailing expansion to an additional sixty schools.

Data from the baseline survey (described in more detail in Section 3.1) can be used to characterize the sample. In Table 1, we compare the characteristics of households in the study sample to state- and country-level averages. In our sample, households include on average seven members, of whom four are children and two are girls. 67% of the sampled households are members of a caste group denoted as OBC, or Other Backward Class; 25% are members of a scheduled caste or scheduled tribe, and the remainder are members of general caste households. Mean land ownership is around six bighas, or approximately one hectare. Study sample households are notably larger, are more likely to be from historically disadvantaged castes/tribes, and have land holdings that are only 25% as high as those of the average Rajasthani household.

This is also a context characterized by relatively low levels of female educational and professional attainment. Among the mothers of the sampled girls, only about 20% reported any post-primary education or engagement in wage employment. More than 90% agreed that a wife should always obey her husband. Among girls in the study sample, 17% were already married when first surveyed, and 84% of girls reported working for pay at that time. Thus, girls in this setting appear to face a number of obstacles to future educational enrollment and academic achievement.

Additional household characteristics, including patterns of income-generating activities, are presented in Appendix Table A1.² Among sample households, 53% reported primary dependence

²Appendix tables are available on the project website: <https://sites.dartmouth.edu/eedmonds/gep/>.

on wage employment, 22% on self-employment in agriculture, 8% on self-employment outside of agriculture, and 16% on casual labor outside of agriculture. Total household consumption in the last month averaged around 26,000 rupees or \$400.

2.2 Intervention

2.2.1 GEP during evaluation period

The Girls' Education Program (GEP) delivered by Room to Read (RtR) is a seven year program that begins in grade six and continues through secondary school. It has two primary goals: encouraging girls to successfully complete secondary school and developing life skills. Since 2007, more than 95,000 girls in nine countries have been enrolled in GEP.

There are two differences between the intervention as analyzed in this evaluation and the broader program. First, the primary data collection for this evaluation was conducted following only two years of program implementation, when the target sampled girls had completed grade 7 rather than grade 12. We will refer to this period (2016–2018) as the evaluation period. Second, the evaluation focuses on a reduced intervention design including only deployment of social mobilizers (“SM”s) who deliver life skills classes and mentoring. The full program additionally includes material support and parent and community engagement, but these program dimensions were not implemented in Ajmer during the evaluation period.

GEP life skills training is delivered in biweekly sessions conducted by SMs during school hours, utilizing a curriculum developed by RtR. In each treatment school, 16 life skills classes were conducted in both grades six and seven. The curriculum is grade-based and emphasizes ten life skills: self-confidence, expressing and managing emotions, empathy, self-control, critical thinking, decision-making, perseverance, communication, relationship building, and creative problem solving. The intervention also focuses on applying these skills to simulations involving time management, education, physical protection and rights, health, and community involvement. It evolves as girls age and regularly revisits topics, adapting to stay age appropriate and relevant. Figures 1 and 2 contain histograms of the number of life skills classes attended by subjects in grade six and seven, respectively. While there was a growth in children who attended no classes between sixth and seventh grade, overall attendance patterns are similar across both grades with complete attendance the mode in both years.

While attending life skills classes, girls may miss some lessons in the primary classroom, though

this varies by school. In some schools, the boys had recess or break time while the girls attended life skills classes; in some schools, normal instruction continued. Any class time lost to life skills sessions would not be expected to significantly impact academic performance given the limited number of annual GEP classes, each of only an hour’s duration, and existing evidence on the low returns to instructional time in Indian schools (see, for instance, Banerjee et al., 2007).

In addition to life skills sessions, the intervention entails biweekly small group mentoring sessions proctored by the SM. In practice, these mentoring sessions served as student-led discussion sections for the life skills lessons taught by the SM. SMs were trained for these mentoring sessions to assist girls in raising more personal difficulties in their lives related to the life skills lessons and to help them to develop more personalized strategies to cope with these difficulties.³ Given GEP’s classroom-based delivery method and the particular emphasis placed on interpersonal skills, including empathy, communication, and relationship building, the program has the potential to impact girls’ social connectedness in addition to strengthening their individual life skills.

40 SMs were employed full-time as a part of this intervention during the two-year evaluation period, with a maximum of 33 employed at any one time.⁴ The typical SM is responsible for two schools (mean of 1.95). GEP aims to have 50 girls per SM. All the SMs (33 years old on average) had completed both secondary and post-secondary education, and all were from Ajmer district; within the district, 58% were from urban areas.

Prior to the launch of the intervention, SMs received 14 days of training, and an additional eight days of training are provided at the start of each subsequent school year. Every eight SMs are supervised by a program assistant, and each SM was observed quarterly to assess the quality of her life skills session and to provide her with support to improve session delivery.

2.2.2 GEP post-evaluation period

In the original design, our partner only had funding for the GEP through grade seven, and thus the endline survey would correspond to the conclusion of the program. However, GEP has always been designed as a project continuing from sixth to twelfth grade, and our partner was successful in attaining programmatic funding to continue the GEP in treated schools beyond the period of

³Earlier GEP descriptions presented these mentoring sessions differently. The description presented here is updated based on the authors’ experiences from talking with students and SMs in Ajmer about their actual experiences in the mentoring sessions. We also note that SMs were referred to as “female role models” in previous project documents, including our published analysis plan.

⁴Of the 40 SMs involved in this evaluation, seven left at some point over the two years and were replaced within three months. One was released for poor performance, and others left for personal reasons such as marriage or migration.

study. They did not begin treating the control schools, and did not extend the program to other, non-study cohorts within treated schools.

However, in 2019 after the evaluation period, RtR added financial support to the GEP in Ajmer. A needs assessment was conducted in fall 2018, and 268 girls began receiving material support in 2019. Material support consisted of in-kind transfers of school supplies, valued at an estimated 500 rupees or approximately \$7.⁵ The addition of material support is not relevant for our primary outcome measures collected in the endline survey, as the survey was conducted prior to the rollout of the material support and the delivery of any associated information (students did not anticipate this future source of material support). However, this change is relevant for supplementary analysis conducted using administrative data, as we also analyze data available post-endline, for eighth and ninth grade.

2.3 Hypotheses

This evaluation examines the impact of GEP on two primary sets of outcomes, school dropout and non-cognitive skills. We also explore additional effects on ancillary outcomes: child marriage, child labor, time allocation and academic achievement. These hypotheses were pre-specified in an analysis plan published prior to the conclusion of the study and available on-line; this plan also included detailed definitions of all variables of interest.

1. Hypothesis 1: GEP has no effect on school dropout in grades six and seven.

The first primary objective of GEP is to increase the completion of secondary school by girls. This objective can be facilitated through the application of the life skills education, the inspiration and support offered by the SM mentor, or the enjoyable dimensions of the program experience. The key outcome measures relevant to this hypothesis include school dropout, progression from one grade to the next, and school attendance. Previous literature suggests that interventions targeting non-cognitive skills, negotiating skills or educational expectations can generate shifts in school enrollment and attendance (Ashraf et al., 2018; Nguyen, 2008; Jensen, 2010). It is also important to note that the relationship between a skills-building intervention and dropout is more plausible in this setting given that financial barriers to enrollment are less salient due to the absence of school fees; both at baseline and

⁵Preparation for assessing need for financial support began after the completion of the endline survey for all respondents other than a small number who had migrated (and was conducted by Room to Read, not the research team).

endline, the modal household pays no school fees for their daughters' attendance.

2. Hypothesis 2: GEP has no effect on life skills.

The second primary objective of GEP is to strengthen girls' life skills. The key outcome measures relevant to this hypothesis include 19 survey-based measures designed to capture life skills and reported by both the child and the parent. In addition, we include four scaled scores for three objective, task-based measures included in the survey. The analysis plan describes the included measures in detail.

Broadly, the questions focus on socio-emotional support, freedom of movement, empowerment, self-esteem, future planning, educational and employment aspirations, marriage expectations and gender norms. We construct normalized indices to capture effects on these domains. In general, previous literature has argued that interventions targeting non-cognitive skills can generate significant effects in adolescence given an overall high level of malleability of non-cognitive skills (Heckman et al., 2006), and evidence from India suggests an intervention targeted at young adults and focused on reshaping gender attitudes had a significant effect in increasing support for gender equality, as well as increasing gender-equitable behaviors (Dhar et al., 2018).

3. Hypothesis 3: GEP has no ancillary effects on child marriage, child labor, time allocation or academic achievement.

Life skills education and mentoring may shift a range of secondary outcomes through direct or indirect channels. The intervention itself may directly induce girls or their families to change their choices around child marriage, child labor or time allocation; in addition, if there are shifts in school progression or non-cognitive skills, these primary effects may generate additional indirect effects in other dimensions of the girl's life, including her level of achievement in school. In general, the literature on effects of interventions of this form on child labor, time allocation and academic achievement is very limited. One recent paper found an intervention targeting equitable gender attitudes among adolescents shifted time allocation for boys, but not for girls (Dhar et al., 2018). Nguyen (2008) finds evidence that an intervention that provides information about future returns to schooling significantly increased test scores in school, but Holmlund and Silva (2014) find that an intervention targeting non-cognitive skills did not enhance cognitive skills. More recently, Ashraf et al. (2018) finds that teaching

negotiation skills to girls in Zambia led to improved test scores and led girls to reallocate time spent on chores across school days.

3 Empirical Strategy

3.1 Evaluation Design

This evaluation is a clustered randomized trial with an allocation rate that was intended to be 1:1, conducted in 119 schools in Ajmer district in Rajasthan between 2015 and 2019. Given that GEP is delivered at the school level to all girls enrolled in the target grades of interest, a cluster randomization is appropriate.

Implementation of GEP was initiated in July 2016 at the beginning of the school year. At the time of design, RtR committed to running GEP in treated schools through the school year ending in the spring of 2018, with the goal of continuing GEP in these schools past that date if possible. Again, for the purposes of this analysis, we define the evaluation period as 2016 to 2018, corresponding to grades six and seven for the treatment girls. As noted above, GEP did in fact continue post-2018 with the addition of material support, and accordingly we also report additional results analyzing administrative data from this post-evaluation period.

3.1.1 Randomization

The selection of schools eligible for inclusion in this evaluation was undertaken between August and November 2015. A team of enumerators visited all schools in Ajmer that included girls enrolled in the relevant grades (six through eight) and collected information about school facilities, staffing, and enrollment. This information was also linked to administrative records about school facilities and enrollment provided by state educational authorities.

The evaluation team and RtR then jointly identified criteria that would determine whether or not a school was eligible for inclusion in the evaluation. These criteria included the requirements that the schools enrolled girls in grades six through eight, did not have any other non-governmental organizations providing life skills curricula to students, and had a classroom in acceptable condition in which a life skills class could take place. The evaluation team then identified the narrowest possible range of enrollments that would yield a sample of schools enrolling 2500 girls in total; the objective was to have a relatively homogeneous sample of schools in terms of size. This yielded the criteria that the school enrolled between 16 and 32 girls in grade five.

Using information collected in the school survey, the research team created a normalized school quality index, composed of measures of teacher experience, teachers’ educational attainment, and classroom and school infrastructure quality. Schools above the median of the index were included in the “high quality” stratum, with the remaining in the “low quality” stratum. School assignment to treatment was conducted separately for the two strata. Randomization was conducted in Stata by the research team, and the list of treatment schools was then communicated to RtR.⁶

In order to identify the target sample of girls, a team of enumerators visited each school between December 2015 and January 2016 to obtain a roster of all girls enrolled in grade five. All female students who were currently enrolled in grade five in these schools as of January 2016 (2,543 female students in total) were eligible for inclusion in the evaluation. There was no further selection of girls within schools.

3.1.2 Data Collection and Processing

The selection of schools and randomization process was conducted in Fall 2015. Baseline data was collected between February and June 2016 prior to the launch of the intervention; a household survey was administered to the child’s caregiver, and a direct interview of the girl potentially eligible for treatment was conducted. All data collection was conducted electronically in SurveyCTO. Details regarding data collection and consent processes are provided in Section A.1 in the Appendix.

Following the baseline survey, the sample girls were revisited for tracking surveys in December 2016 and December 2017. The endline survey was conducted between July 2018 and December 2018. Baseline and endline surveys included both a household module and a child module for every girl in the sample, while only girls were surveyed for each of the two shorter tracking surveys. Consent was obtained separately for each subsequent survey.

3.1.3 Administrative and Qualitative Data Collection

In addition, administrative records from schools and RtR were also collected throughout the evaluation period and up to eight months post-endline (through July 2019). Administrative data from

⁶Following the initiation of data collection, it was discovered that three of the schools selected to be in the sample in fact did not enroll girls past grade five; for the upper-level grades, these were single-sex schools including only boys. During the sample selection process, these schools were incorrectly designated as including higher-grade girls as well. These three schools (two treatment and one control school) were dropped, and an additional three schools were selected to replace them, constituting an additional third strata. The replacement process for these schools entailed identifying 12 schools that met the eligibility criteria if the enrollment window was slightly lowered to 15. Three schools were randomly chosen to join the sample among the 12, and of these, two were randomly assigned to the treatment group.

schools include enrollment and reported grades. It should be noted that in the majority of cases, our test score records are drawn from gradebooks where teachers hand copied student marks; accordingly, the potential for measurement error is non-trivial, a challenge we explore further in the results presented in Section 4.

Two points should be noted about the administrative data collection. First, given that its availability was not foreseen *ex ante*, this analysis was not pre-specified. Second, given the post-endline data collection, administrative data is available beyond the core evaluation period: grades are available through the conclusion of grade eight, and dropout information is available through early in grade nine.

In addition, administrative data from RtR report girls' participation in the intervention, including life skills sessions and mentoring. The research team also oversaw qualitative data collection at each phase of the evaluation, including in-depth interviews with girls in a subset of treatment schools as well as their caregivers.

3.1.4 Evaluation Sample

At baseline, the survey team visited every one of the 2,543 girls on the enrollment lists provided by sampled schools; the survey was conducted before students or their families were informed about the life skills education program. Ultimately, any girl on the enrollment lists with either a completed household or child survey is considered to be enrolled into the evaluation. (In some cases, there are multiple sample girls in the same household.) Out of the 2,543 female students on the grade five enrollment lists, a total of 2,459 girls from 2,382 households were enrolled into the evaluation sample. Thus the evaluation includes 97% of the girls in the sampled school rosters.

However, not every girl who was part of the evaluation sample was interviewed at baseline; there were cases in which only the household survey was completed, as well as a smaller number of cases in which only the girl survey was conducted.⁷ There were 2,353 household surveys conducted at baseline, which provide parent-reported data for 2,427 girls, and 2,399 individual girl surveys conducted at baseline. A flow chart summarizing the sample of girls surveyed and their inclusion in different evaluation phases can be found in Figure 4. There were 84 children who were on the school enrollment lists but excluded from the evaluation because of failure to complete any component of

⁷In addition, 16 girls living in 14 separate households from one primary school were omitted from the baseline in error. A different set of students enrolled in an alternative, adjacent primary school that is outside the evaluation sample were surveyed in their place. Given that these girls were not intended for inclusion in the sample, their data was subsequently dropped, and the correct set of girls were surveyed from the first tracking survey forward. These girls are considered to be enrolled in the evaluation, though they were not surveyed at baseline.

the baseline survey.⁸

Average summary statistics for the child outcomes mapped out in the evaluation design plan are provided in Table A2. 3% of the sampled children had already dropped out of school at baseline. However, conditional on enrollment, 89% attended school in the past week, and girls attended 79% of the days the school was open. While the modal girl in our study is at an appropriate age for grade 5 (age 10-11), approximately a third of the sampled girls were older than that. 17% of the sampled girls were married at baseline. 87% of respondents engaged in child labor during the twelve months before the baseline survey, and 64% of respondents engaged in hazardous child labor. On average children were not working full time.

Attrition The analysis sample includes all sampled girls who were represented in the endline survey, conducted between July 2018 and December 2018. At endline, 2,387 child surveys and 2,358 household surveys were conducted. There were 48 girls (in 47 households) for whom a household survey was conducted without a child survey.⁹ There was also one girl surveyed whose household did not complete an endline survey; in this case, the head of household consented for the girl’s participation but declined to complete the household survey. In addition, 24 girls in 23 different households attrited fully at endline with no data collection completed.¹⁰ Attrition patterns are summarized graphically in Figure 5.

If we examine patterns of attrition with respect to treatment arm, we observe the following. Among the 24 girls who fully attrited, 15 are from schools assigned to the control arm and 9 are from schools assigned to the treatment arm. These figures correspond to attrition rates of 0.7% and 1.2% in the treatment and control groups, respectively; the probability of full attrition is not significantly correlated with treatment, conditional on strata fixed effects ($\beta=-.004$, $p=.362$). Among the 72 girls who attrited from the girl survey, 45 are from schools assigned to the control

⁸Of these 84 cases, 34 were from households that had permanently migrated prior to the date on which the survey team visited the community — a fact reported by neighbors or other community informants — or simply could not be located. 33 were excluded because they did not provide consent. The reasons for non-inclusion for the remaining girls varied but included illness or death of the child (4); parents who were uniformly unavailable during survey hours and thus could not be surveyed or provide consent for the child to be surveyed (3); and cases in which the child was away from home and parents declined to participate in her absence (10).

⁹In 14 cases, consent was declined for the girl survey. 19 girls had migrated away from their households permanently; two had migrated temporarily and had not returned by the point at which the survey concluded. Four child surveys were not completed due to the death of the child, and nine child surveys were not completed due to child disability. In these nine cases, the child was similarly not surveyed at baseline, but a household survey was completed at both baseline and endline.

¹⁰In 10 cases, the household had migrated and could not be reached for follow-up. Consent was declined in 12 cases. In one case, a partial survey was completed but the household declined to continue due to limited time, and in one case, no information was available about the household’s whereabouts.

arm, and 27 are from schools assigned to the treatment arm. These counts imply attrition rates of 2.2% and 3.6% in the treatment and control groups, respectively; this difference is significant at the 10 percent level ($\beta=-.014$, $p=.087$). In Appendix Tables A9 through A13, we conduct a series of bounding exercises to verify that differential attrition on the girl survey does not substantively affect the primary research conclusions under reasonable assumptions regarding the distribution of missing values.

3.2 Statistical Model

To identify the impacts of the intervention, we employ three specifications. First, we estimate an ordinary least squares (OLS) regression of each outcome of interest on an indicator variable for treatment assignment and indicator variables for randomization strata. No additional baseline control variables are included. The equation of interest can thus be written as follows for each child outcome, denoted Y_{ist} for child i in school s measured at time t . T_s denotes the dummy for treatment assignment, and μ_s denotes the randomization strata fixed effect for school s .

$$Y_{ist} = \beta_1 T_s + \mu_s + \epsilon_{ist} \quad (1)$$

Second, we estimate a specification including baseline control variables in which the outcomes of interest are regressed on an indicator variable for treatment assignment, indicator variables for randomization strata, a vector of age dummies, a vector of dummies capturing the most important type of employment in the household at baseline, and a control variable that measures the lagged (baseline) value of the relevant outcome. The equation of interest can thus be written as follows:

$$Y_{ist} = \beta_1 T_s + \beta_2 Y_{is,t-1} + \mu_s + \gamma_i + \lambda_i + \epsilon_{ist} \quad (2)$$

$Y_{is,t-1}$ denotes the baseline value of the outcome, γ_i denotes a vector of age dummies, and λ_i denotes a vector of indicator variables for the most important type of employment in the household at baseline.¹¹

¹¹For the family of outcomes corresponding to school progression and completion (Hypothesis 1), age at enrollment and maternal education are additionally included as control variables in equation (2); this methodology was pre-specified in the analysis plan. If the baseline control variable is missing because either the household or child survey was not conducted for a particular girl at baseline, the missing value is coded as zero. Additional dummy variables equal to one for observations with missing values are included for each baseline covariate. For non-cognitive outcome measures added at endline (the Rotter locus of control, the perceived stress index, the Rosenberg self-esteem index), we control for lagged values of overall life skills indices. For the ASER test scores added at endline, we control for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school and grades in grade five as reported in administrative data. These methodologies for addressing missing baseline values were all

Third, we estimate a specification that adds to specification (2) additional controls for baseline variables where imbalance was reported between the treatment and control households in Tables A3 through A5 (discussed in more detail below). These baseline variables are denoted $\xi_{is,t-1}$.

$$Y_{ist} = \beta_1 T_s + \beta_2 Y_{is,t-1} + \beta_3 \xi_{is,t-1} + \mu_s + \gamma_i + \lambda_i + \epsilon_{ist} \quad (3)$$

In all specifications, standard errors will be clustered at the school level. Our sample includes 119 clusters. There are a large volume of hypotheses tested regarding life skills. There will be false discoveries (type 1 errors). For all life skill measures, we present false discovery rate adjusted q-values computed across all life skill outcomes using the same specification (Benjamini and Hochberg, 1995).

3.2.1 Balance

We evaluate balance at baseline across 61 variables relevant to our analysis in Tables A3 through A5. In the column labeled “Difference”, we report the coefficients from a simple regression in which the characteristics of interest are regressed on a treatment indicator and strata dummies, clustering standard errors at the school level. With 61 tests, we expect type 1 errors. We also present false discovery rate adjusted q-values using the same specification (Benjamini and Hochberg, 1995).

For household characteristics, we reject the null hypothesis of no difference with treatment for only one of the variables in Table A3 (treated girls are more likely to be from Other Backward Class households). We also estimate a seemingly unrelated regression specification that tests the joint null hypothesis that the treatment coefficient is equal to zero, and fail to reject the null ($p = .138$).

For child characteristics, we reject the null of no difference (at the 10 percent level) for 7 of the 42 characteristics examined. The p-value on a joint test of significance of the treatment coefficient across all equations, constructed as described above, is 0.003, implying that we reject the hypothesis that child characteristics do not differ systematically based on treatment status. 6 of these 7 have a false discovery rate above 50 percent. Only the indicator for whether child works has a p-value below 0.10 and a false discovery rate below 25 percent (it is 23 percent). Given the high false discovery rates, we do not think these differences in Tables A3 through A5 are a concern, but in all results below, we include a specification that controls for all of the variables with individual differences significant at 10 percent as a robustness check. In practice, these controls add precision,

pre-specified in the evaluation design plan addendum.

and we never reject the null hypothesis of no change in treatment effects with the added controls.

4 Findings

4.1 Dropout

To analyze the first hypothesis, we report the results for outcome variables corresponding to academic progression in Table 2. In all tables reporting empirical results, Panel A reports the results of estimating equation (1) without additional control variables, Panel B reports the results of estimating equation (2) controlling for baseline values of the outcomes of interest, and Panel C reports the results of estimating equation (3) controlling for variables imbalanced at baseline. In this table, Columns (1) to (4) report regressions using data drawn from our endline survey; Columns (5) through (8) report regressions estimated using administrative data.

The estimates in Column (1) of Panel A suggest that GEP reduced dropout by 3.3 percentage points, or 25% relative to a mean dropout rate of 13.2% in the control group. The estimates in Panel B and C are of comparable magnitude, and the coefficient is significant at the 5 percent level in Panel C. In Column (2), we observe a corresponding increase in grade progression of very similar magnitude. The fact that the estimated coefficients for dropout and grade progression are nearly symmetric suggests the effects of the program operate solely through reducing dropout, and not via shifts in the probability of promotion to a new grade conditional on enrolling in school.

However, in Columns (3) and (4), there is no statistically significant impact of the intervention on attendance conditional on enrollment, reported using a binary or a continuous variable. (298 observations are missing for these endline attendance measures because of temporary school closures; in addition, attendance is not reported for girls who dropped out.)¹²

In Columns (5) through (8), we utilize administrative data to further analyze dropout in grades six through nine, particularly in the transition to high school; ninth grade is the first year of high school in this context. We estimate the same specifications, utilizing dropout as reported in grade five as the baseline control variable in Panels B and C. In interpreting these results, it is important to note that grades eight and nine correspond to the post-evaluation period during which some limited material support was rolled out. Moreover, as explained earlier, analyses based on administrative data were not pre-specified.

¹²We also examine effects on attendance using administrative data reported by the schools, and similarly observe a null effect; these results are not reported, given that we conjecture measurement error is substantial.

In Column (5), we observe the estimated effect of treatment on dropout in grade six (after one year of intervention) is already negative, though the estimate is not statistically significant. Columns (6) through (8) report negative effects of the treatment on dropout that are monotonically increasing in magnitude, and significant for grades eight and nine. (Since the endline survey was conducted at the start of eighth grade, this estimate captures dropout at a point intermediate between the conclusion of grade seven, captured in the data reported in Column (6), and at the conclusion of grade eight, captured in Column (7).)

It is important to note that because the grade nine records were collected at the start of the school year in late July 2019, they differ from the other administrative records that are based on attendance throughout the school year. For grade nine, data is available only about how many days the child attended school in the last seven days, and how many days the school was open in that period. Hence, defining dropout based on this measure will have substantive measurement error if there is any irregularity in attendance, and the evidence suggests irregularity is non-trivial. Accordingly, the findings for grade nine are estimated using a modified version of specification (1) that also controls directly for the number of days the school was reported open. Without controlling for this form of measurement error from days open, we find a noisy zero effect of treatment on dropout in grade nine (not presented).

To summarize these results, Figure 6 presents the coefficients estimated in Panel A in graphical form. For each grade, we observe the cumulative effect on dropout to be increasing, although we cannot reject the null that the impact of treatment on dropout is constant from grades six to nine. Given rising dropout across grades in the control sample, the percentage change in dropout is relatively constant. It is impossible to identify whether the results reported for administrative data in grade eight and nine reflect any impact of the needs assessment or the subsequent material support. In Appendix Table A6, we estimate the degree of potential bias due to missing data in these school dropout analyses that rely on administrative data; these bounding exercises support the conclusion that the impact of GEP on dropout continues into high school.

4.2 Life Skills

To analyze the second hypothesis, we evaluate treatment effects for 19 indices built from survey questions as well as four life skills measures built from demonstration tasks. All index construction was pre-specified. For all life skills measures, we again report the results of estimating specifications (1) through (3) in Panels A through C of Tables 3 through 5, respectively. For outcome variables

that are dummy variables or measures of time and effort, we report the magnitudes in terms of percentage effects. For outcome variables that are calculated as indices, however, we follow conventions in the literature (and our analysis plan) to report the effect magnitudes in terms of standard deviations in the control arm.

4.2.1 Child Measures

We find substantial and statistically significant enhancements in life skills for girls assigned to treatment. Coefficients for ten life skills indices constructed based on girls' responses are reported in Table 3. GEP assignment increases the index of socio-emotional support by 0.07 standard deviations, increases the empowerment index by 0.09 standard deviations, increases the index of future planning by .07 standard deviations, and increases the gender norms index by 0.09 standard deviations. These effects are consistent in magnitude and significance across specifications and have associated Q-statistics that are uniformly smaller than 0.25 (typically between 0.01 and 0.10). In addition, there is some evidence of a positive effect of the treatment on the enumerator assessment of the girl, an increase of .07 to .10 standard deviations that is significant in Panels B and C.

There are, however, no statistically significant effects on the freedom of movement index, the educational and employment aspirations index, or Cantril's ladder. There is a statistically significant decline in the marital expectations index of .3 standard deviations.¹³ The shift in the marital expectations index is driven primarily by a shift downward in the top of the distribution of desired and expected marriage ages. The GEP curriculum strongly emphasizes 18 as the appropriate minimum age of marriage, and there is some evidence that treatment girls are then more likely to report 18 as the desired age of marriage relative to both younger and older desired ages. Our findings of significant enhancements in life skills highlight a potential pathway to explain the measured declines in dropout associated with program participation. We return to the discussion of mechanisms in Section 5.

In addition, evidence suggests that these effects do not simply reflect a process in which girls repeat certain socially desirable responses that were explicitly taught in the GEP curriculum. In a supplementary analysis, we classify 90 individual life skills questions posed in the endline survey based on whether they are explicitly addressed in the grade six or grade seven curriculum or addressed only indirectly. It is important to bear in mind that the distinction between whether an

¹³Seven observations are missing for the measures reported in this table, corresponding to the seven cases in which the respondent elected only to respond to the first section of the child survey.

item is explicitly addressed is based on whether we could identify an exact curricular match to the question; all of the life skills that we classify as not explicitly addressed may be indirectly addressed in the program. We then construct separate indices characterizing responses to explicitly versus indirectly addressed questions, and estimate the treatment effects for these indices.

We find a treatment effect of 0.056 SD for the explicitly addressed questions index (standard error of 0.018) and an effect of 0.016 SD for the indirectly addressed questions index (standard error of 0.019). However, we know that the marital expectations questions show an effect in the opposite direction of that hypothesized, and these questions are primarily in the indirect questions index. If we exclude the five questions related to marriage age that are included in the marital expectations index, the treatment effect for the indirectly addressed questions index rises to 0.040 SD (standard error of 0.20), and we cannot reject the hypothesis that the treatment effects for indirect and direct questions are equal in magnitude. Given this evidence, we argue that the observed effects on life skills do not reflect merely parroting back the curriculum to the enumerator. It is also worth emphasizing that at the point of the endline survey, the majority of the subjects had not had a life skills class for six months, as the intervention concluded before seventh grade exams and had not yet re-commenced in the new school year.

Coefficients for the four demonstration task measures are reported in Table 4. Here, there are no statistically significant treatment effects on the associated outcomes; we cannot reject that girls in treatment schools perform the same as girls in control schools on the delay discounting, mirror drawing, and scavenger hunt tasks. In general, the magnitudes of the estimated coefficients are small and consistent across panels.¹⁴ Based on reports from the field staff and our own observations, measurement challenges seem likely to explain these null findings.¹⁵ Prior to endline, we opted to add previously validated measures of psycho-social well-being to the survey, and amended the analysis plan prior to data collection to include the Rotter index of locus of control, the perceived stress index, and the Rosenberg self-esteem index. The results estimated using these added psycho-social well-being outcomes are reported in Columns (5) through (7), and we fail to reject null

¹⁴In total, seven observations are missing from the analysis for both the future discounting and scavenger hunt measures, corresponding to the seven cases in which the respondent elected only to respond to the first section of the child survey. 70 observations are missing for time spent on mirror drawing measure, corresponding to the 70 respondents who did not attempt any mirror drawings.

¹⁵Specifically, we observed that girls' efforts on the mirror drawing task varied based on the particular environment in which they were surveyed. Indeed, the within-girl correlation between baseline and endline measures is only 0.06. For the scavenger hunt, variation in the time between the two required surveyor visits (one to introduce the scavenger hunt and one to assess scavenger hunt success), in addition to overall delays in the timing of surveyor re-visits, seems to have limited the signal value of the associated measures.

effects.¹⁶

4.2.2 Parental Reports

The results for parental reports of life skills reported in Table 5 suggest that the treatment did not significantly shift parents' assessment of their daughter's life skills. The estimates are generally imprecise, heterogeneous in sign, and small in magnitude across alternative specifications.¹⁷ The one dimension along which we find significant effects is the index of parental perceptions of girl's strengths, and here the coefficient of interest is negative. This negative parental perception is discussed in detail in Section 5; essentially, parents perceive treated girls as more selfish.

4.3 Ancillary Outcomes

The effects of treatment on ancillary outcomes are reported in Table 6 for child marriage and child labor, and in Table 7 for time allocation, scores on researcher-administered tests (ASER test scores), and scores on year-end exams in mathematics, Hindi and English administered in school. Again, analysis using the administrative data as presented in Columns (6) through (8) of Table 7 was not pre-specified. However, we employ the same specifications utilized in the primary analysis, utilizing the GPA reported in fifth grade as a baseline value.¹⁸

In general, the effects of treatment on ancillary outcomes are insignificant and inconsistent in sign. This is true for child labor both on the intensive and extensive margin, for time allocated to school and ASER test scores (reported for all girls surveyed at endline), and for administrative test scores in grades six, seven and eight (reported for all girls observed in school-reported data in each year).¹⁹ For grade six, there are negative effects of treatment on administrative test scores that are larger in magnitude and statistically significant when controls are included; this is driven entirely by the mathematics score, and the magnitude is not inconsistent with the other estimates in the

¹⁶Seven observations are missing for each of these endline measures, corresponding to the seven cases in which the respondent elected only to respond to the first section of the child survey.

¹⁷Four observations are missing for parental perception of girl's self-efficacy, as the parent answered "Don't know" to all the relevant questions.

¹⁸There is heterogeneity within schools across years and across schools in how year-end exams are scored. For comparability purposes, we have computed an average GPA across the three tests based on the letter grades associated with the numerical test scores. Note that the grade eight test results also correspond to the post-evaluation period in which material support was rolled out.

¹⁹In Table 6, one observation is missing for the majority of the outcomes reported in the table, corresponding to one respondent who did not answer question 311 in the child survey. In Table 7, controls for baseline outcome values cannot be included for cognitive test measures since cognitive tests were not conducted at baseline; specification (3) instead includes controls for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school, and grades as reported in grade five.

table. By grade eight, the magnitudes of the negative coefficients have diminished, and the point estimates are consistent with small positive or small negative effects of treatment on test scores.

In addition, the estimated magnitude of the coefficients estimated using ASER test scores is broadly consistent with the estimated magnitude of the coefficients estimated using administrative test score data.²⁰ Further exploration of the effects on school test scores, including analysis of any potential bias induced by selection into test attendance, is reported in Appendix Tables A7 and A8. In sum, we find little evidence that selection into test attendance can explain the limited test score impacts that we identify.

4.4 Additional results

In Appendices A.2 and A.3, we further explore the robustness of the primary results to alternative assumptions about attrition, and find that the primary effects remain generally consistent in magnitude and significance. In Appendix A.4, we examine heterogeneous effects with respect to a number of pre-specified baseline covariates, and find no evidence of meaningful heterogeneity.

5 Discussion

Our principal experimental findings suggest that GEP reduces dropout and improves the expression of life skills among adolescent girls after two years of life skills education, and this section aims to understand the channels for these two results. All previous sections of this paper (except for the administrative data analyses presented in Sections 4.1 and 4.3) were pre-specified. This analysis, however, should be considered exploratory, and it is motivated in part by findings from our qualitative field work.²¹

In a conventional treatment of agency within economics, household decisions result from an aggregation of parent and child preferences, and improvements in agency are modeled as a relative

²⁰The endline survey was conducted approximately four months after the seventh grade exams. Hence, we would expect the seventh grade results in Column (7) to be most comparable to the results in Columns (3)-(5). The most comparable exercise would be to average across the three ASER scores. For those three scores, we find treatment is associated with a statistically insignificant -0.02 reduction in test scores. Compared to a mean score of 2.58, this reduction rounds up to a one percent reduction in ASER scores. The 95 percent confidence interval for the seventh grade results in Panel A of Column (7) range between a 13 percent decline in test scores and a 2.5 percent increase. Hence, the seventh grade scores on the in-school tests are consistent with what we are finding in the survey-administered tests.

²¹There are two qualitative data collection efforts related to this project. First, there was a formal qualitative evaluation led by Joan DeJaeghere, employing a pre-determined research design and involving formal, structured interviews (DeJaeghere et al., 2018). Second, the authors of this paper conducted less-structured qualitative interviews with available subjects, teachers, and SMs during initial piloting, during baseline surveying, and after endline surveying.

shift in the weight received by the child’s preferences. Hence, our dropout findings could be understood as demonstrating that life skills education enhances the ability of girls to negotiate with their parents in order to remain in school. We argue, however, that the reductions in dropout we observe do not primarily stem from a shift in the intrahousehold bargaining process. Rather, the evidence is consistent with the hypothesis that the intervention raises the value of child time in school as perceived by the child. This increased value could in turn reflect two channels: stronger social relationships that yield higher utility returns to time in school, or an increase in the perceived returns to schooling. In general, the evidence is more consistent with the former channel.

There is ample evidence in our findings to suggest that treatment improves girls’ ability to advocate for themselves. In our qualitative work, teachers, SMs, and girls all highlighted that the intervention led to perceived growth in girls’ self-confidence and willingness to advocate for themselves. We see this in the data, as evident in the improvement in the empowerment and self-esteem indices reported in Table 3 above. Individual survey question responses provide additional support. In Table 8, we replicate our base empirical specification (1) to examine a number of specific questions. In Panel A, we observe an increase in the probability girls state that they alone decide if they will go to school, and an increase in the probability girls feel they are solely responsible for deciding if they continue schooling past the end of primary school. The relatively low cost of schooling in our study setting may amplify changes in perceived autonomy by reducing financial dependence on parents. At endline, the modal household does not pay any school fees for their daughter, and additional costs (uniforms, books, etc.) represent only five percent of total household expenditures. That said, treatment affects autonomy along other dimensions as well: we observe significant increases in the probability that girls discuss marriage with their parents, and the probability girls think they alone will select their future profession. The magnitude of all of these effects is between 15 and 25 percent.

Parents’ survey responses are also consistent with this growth in child assertiveness. We have already seen in Table 5 that treatment leads to more negative parental perceptions of girls’ strengths. Essentially, parents see treated girls as more selfish. This finding is consistent with the evidence in Table 8 that treatment girls assert more control over their lives; to parents of teenagers, this can be annoying. Additional evidence on parental perceptions of girls’ attitudes and behaviors is presented in Panel B of Table 8. Treatment girls are perceived to be less willing to help, less considerate and, while not statistically significant, are seen as less likely to honor the requests of adults.

Overall, we view these findings as a reflection of girls standing up for themselves and being

more assertive about their own interests. However, this growth in self-advocacy may play a less central role in explaining dropout findings than one would expect in the classical agency model; more specifically, it is not clear that girls' preference for more education significantly diverges from parental preferences. In the control sample at endline, two-thirds of parents report that they expect their daughters will complete at least senior secondary school, and 71 percent expect their daughters will complete at least five more years of schooling (beyond grade seven). Moreover, 54 percent of parents think that being well-educated is an important characteristic for a potential daughter-in-law, and 74 percent believe that girls need to pursue higher education. As such, it is not obvious that parents are less enthusiastic about their daughters' schooling than the girls themselves. These patterns are also consistent with Bursztn and Coffman (2012), who find that parents value their children's attendance at school, and accordingly value conditionality in cash transfers as a strategy to manipulate child school attendance.

In addition, we do not find any statistically significant impact of treatment on parents' attitudes towards girls' schooling or on parents' perceptions of parent-daughter communication, and we see no evidence of the time use or academic achievement effects found in past work that has sought to isolate the downstream impacts of changing within-household negotiating dynamics (Ashraf et al., 2018). Thus, while it is certainly plausible that changes in the weight placed on child preferences contribute to reduced dropout, the body of evidence indicates that other mechanisms play a more central role. More specifically, we argue that the intervention appears to change the marginal utility associated with time spent in school and thereby raise the opportunity cost of dropout. This increase in the opportunity cost of dropout could reflect increased social support and social engagement in schooling, and/or an increase in the perceived returns to education. In this case, the evidence is stronger for the former channel.

Our qualitative work emphasizes the importance of stronger social support in raising the shadow value of school attendance; when interviewed, girls emphasized that GEP participation resulted in stronger friendships. Girls' reported enjoyment of the life skills classes themselves would also be expected to raise the shadow value of continued enrollment, as the classes are conducted during the school day. Though strengthening social supports is not an explicit goal of GEP, the curricular focus on interpersonal skills, including empathy, communication, and relationship building, suggests that the program is likely to influence girls' social connectedness in addition to strengthening their individual life skills. The qualitative evidence is reinforced by quantitative findings related to the effect of treatment on social engagement, as measured by parental survey responses, child responses,

and our time use survey of children. Girls in treatment schools are significantly less likely to be identified as preferring to be alone; they are more likely to meet friends out of school, more likely to have a place to meet female friends, and more apt to report they have a place to stay if they needed one. These findings are reported in Panel A of Table 9, and we find a sizable (0.10 SD) treatment effect on an index that aggregates across these separate social engagement measures.

Consistent with their increased desire to socialize, treatment girls appear to be spending more time in contact with friends. Panel B of Table 9 demonstrates that this increased social engagement results in treated girls reporting longer travel times to and from school (they are not differentially likely to report changing schools relative to the control group and this finding holds conditional on continued enrollment). Control students report that they spent 37 minutes going to and from school on a typical day in the last week, while treatment girls spent an extra nine minutes. Treatment girls also spend more time on their mobile phones. They report an average increase in mobile use of 0.3 minutes during a typical day in the last week. While this change is small, it is 15 times greater than baseline usage in the control group. Overall, few girls report using a mobile, but for those that do, the observed increase in usage corresponds to an additional 37 minutes of mobile use in a day.²² Total social time increases from 39 minutes a day in the control group to 49 minutes in the treatment group, a 25 percent increase.

Of course, the growth in social supports need not impact dropout solely by changing girls' desire to spend more time with friends. Improved social supports may also help girls to overcome salient socioemotional challenges. In our qualitative work, we heard frequently that teasing is a significant challenge; persistent teasing increases dropout, either because girls seek to avoid teasing or due to pressure from parents who are concerned about extensive girl - boy interactions. While girls learn in one life skills class to ignore teasing, one girl ignoring teasing may or may not work in practice. In contrast, an entire class of girls ignoring such teasing may prove much more effective, suggesting that strengthened friendship networks may serve to amplify direct effects of treatment in addition to independently influencing dropout propensity.

While treatment clearly fosters social interaction and associated benefits, a second channel through which the treatment could increase girls' utility from schooling and reduce dropout is by increasing the perceived future returns to education itself. Here, however, the data provide more limited support, as illustrated by Panel C of Table 9. Column (1) examines as an outcome a variable

²²The growth in time on the phone does not represent growth in access. 87 percent of our control group has access to a mobile (either they own one themselves or their family does and they can access it) and treated girls are only one percentage point more likely to have access (with a standard error of 1.5 percentage points).

that takes on the value of one if the respondent girl aspires to work in an occupation that requires completing higher secondary schooling. Treatment does not significantly impact this outcome. A respondent's intention to work for pay is often employed as a proxy for perceived future returns to education, and we also do not see a substantive change in this outcome associated with treatment. 76 percent of respondents want to work for pay in our control group, and treatment raises this desire by only one percentage point. Finally, we asked girls directly how much education they would like to complete. 93 percent of control group girls wanted to complete secondary schooling, and the confidence interval on the treatment effect ranges between -2.5 and 2.6 percentage points.

While we did not ask girls about perceived future returns to education due to challenges that arose in piloting related questions, we did ask whether girls felt that they should not get educated because they will get married. 22 percent of control respondents replied that girls should not get educated. Treatment reduces this rate by 5 percentage points, or 21 percent. This difference might reflect a change in the perceived value of education. Hence, while we do not view as compelling the evidence that treatment raises perceptions of the value of education, we are unable to exclude this channel.

Overall, we interpret our results as consistent with the hypothesis that increased social engagement is an important channel for the observed decline in dropout. It does not seem that our findings are consistent with an increase in the weight placed on the child's preferences. These findings related to increased social engagement also help to resolve a puzzle: we find that treatment reduces dropout without improving attendance or school performance. An increase in social support may make it easier to miss school, as friends can assist to make up any missed material. Hence, girls want to go to school more, but missing school becomes less costly; the net effect is ambiguous. For academic performance, additional social engagement (including spending more time on the phone) may offset whatever gains might otherwise result from increased future planning and growth in motivation. We also observe that treatment leads to a decline in private tutoring; this is presumably replaced by group-based, social activities, as overall our primary results do not suggest there is any substantive change in time spent studying.²³

²³One obvious concern might be that negative selection in students who do not dropout is masking the gain in test scores for higher-achieving students. While dropouts are indeed negatively selected, we do not see improvements among students who were higher-achieving at baseline.

6 Conclusion

In this paper, we analyze evidence around an intervention aimed at improving the life skills of adolescent girls, targeted to a sample of girls enrolled in sixth and seven grade in Rajasthan, India. Evidence suggests that over a two-year follow-up period, the intervention was successful in reducing dropout and enhancing girls' non-cognitive skills over a range of dimensions linked to agency, social support, and goal-setting. However, there were no statistically significant impacts on school performance, attendance or time allocated to school. We interpret these results as evidence that even in a setting of gender disadvantage, adolescent girls can influence their own schooling. The intervention increases girls' social support in school — thus increasing their desire to enroll — and perhaps helps them to advocate for their own needs within the household.

The finding of a lack of improvement in test scores is consistent with many other studies in the life skills space (e.g. Holmlund and Silva, 2014; Delavallade et al., 2017). In our context of reduced dropout, this naturally raises the question of whether continued school enrollment is valuable in itself. In fact, one principal we met in our qualitative work raised this exact question herself based on her sense (not our findings) that girls were continuing in school without doing any better in school. It is obviously possible to have financial returns to education that are not easily measurable with tests in English, Hindi and Mathematics. It is also worth remembering that the education results are compared to the control population. Eighth graders still know more than seventh graders, even if that does not change differentially with treatment. Continued enrollment may also facilitate delayed fertility, and this may be especially relevant in our setting with pervasive child marriage. The literature on returns to female education further highlights the value of staying in school for the life skills, experiences, and social relationships that education can help foster. It is plausible that our findings reflect feedback between social relationships and schooling: girls stay in school because of stronger social relationships and continued schooling strengthens those relationships, which may be important later in life regardless of whether there is meaningful learning.

Room to Read's Girls Education Program usually includes outreach and financial support that was not delivered in the study area. While we are able to evaluate the impact of the life skills component of the program in isolation, our study is nonetheless related to recent research evaluating multifaceted interventions that combine life skills training with other social services such as Save the Children's Safe Spaces (Buchmann et al., 2017) and BRAC's Empowerment and Livelihood for Adolescents (ELA) program (Bandiera et al., 2019b,a). While both Safe Spaces and ELA target

older girls than our study, they both document improvements in schooling. Our finding that life skills training (separate from the other components of ELA or Safe Spaces) increases education in part through building social relationships highlights the potential importance of that specific component of these multifaceted programs. Relatedly, our finding that life skills alone is not sufficient to influence some of the important life decisions that ELA impacts also highlights the additional value of other components of the multifaceted approach, despite the non-experimental evidence in Bandiera et al. (2019b) that emphasizes the contribution of life skills training in particular.

Our study delivery method also highlights the tradeoffs inherent in targeting decisions related to life skills programs. Both ELA and Safe Spaces uses time and space outside of schools. This allows them to reach more marginalized girls not associated with a school absent the program. However, they face much lower take-up than our school based intervention. While 85 percent of our subjects are still engaged after two years, Safe Spaces only managed to induce 56 percent of girls to attend one class, and ELA take-up in Uganda is below 25 percent. In our context, not only is it easy to reach girls within school, but the girls already have within-school social relationships that can be leveraged. To the extent that the reinforcement and deepening of those social relationships drive our dropout results, such dynamics might not be present in an intervention targeting out of school girls.

7 References

- Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn, “Negotiating a better future: How interpersonal skills facilitate inter-generational investment,” 2018.
- Baird, Sarah J, Ephraim Chirwa, Jacobus De Hoop, and Berk Özler, “Girl power: cash transfers and adolescent welfare. Evidence from a cluster-randomized experiment in Malawi,” Technical Report, National Bureau of Economic Research 2013.
- Bandiera, Oriana, Niklas Buehren, Markus P Goldstein, Imran Rasul, and Andrea Smurra, *The Economic Lives of Young Women in the Time of Ebola: Lessons from an Empowerment Program*, The World Bank, 2019.
- , —, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, “Women’s empowerment in action: evidence from a randomized control trial in Africa,” *American Economic Journal: Applied Economics*, 2019, *forthcoming*.
- Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden, “Remedying Education: Evidence from Two Randomized Experiments in India,” *Quarterly Journal of Economics*, 2007, 122 (7), 1235–1264.
- Becker, Gary S, “Investment in human capital: A theoretical analysis,” *Journal of political economy*, 1962, 70 (5, Part 2), 9–49.

- Benjamini, Yoav and Yosef Hochberg**, “Controlling the false discovery rate: A practical and powerful approach to multiple testing,” *Journal of Human Capital*, 1995, 57, 289–300.
- Berry, James**, “Child Control in Education Decisions An Evaluation of Targeted Incentives to Learn in India,” *Journal of Human Resources*, 2015, 50 (4), 1051–1080.
- Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, and Iman Sen**, “Power vs money: Alternative approaches to reducing child marriage in Bangladesh, a randomized control trial,” *Unpublished Manuscript*, 2017.
- Bursztyn, Leonardo and Lucas C Coffman**, “The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas,” *Journal of Political Economy*, 2012, 120 (3), 359–397.
- , **Georgy Egorov, and Robert Jensen**, “Cool to be smart or smart to be cool? Understanding peer pressure in education,” 2017.
- Cunha, Flavio, James J Heckman, Lance Lochner, and Dimitriy V Masterov**, “Interpreting the evidence on life cycle skill formation,” *Handbook of the Economics of Education*, 2006, 1, 697–812.
- DeJaeghere, Joan, Aditi Arur, and Devleena Chatterji**, “Qualitative Longitudinal Study Endline Report,” 2018.
- Delavallade, C, A Griffith, and R Thornton**, “Network partitioning and social exclusion under different selection regimes,” *Unpublished, International Food Policy Research Institute, Washington, DC*, 2016.
- Delavallade, Clara, Alan Griffith, Gaurav Shukla, and Rebecca Thornton**, “Participation, learning, and equity in education: Can we have it all?,” 2017.
- Deming, David J**, “The growing importance of social skills in the labor market,” *The Quarterly Journal of Economics*, 2017, 132 (4), 1593–1640.
- Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch**, “Comparative cost-effectiveness analysis to inform policy in developing countries: A general framework with applications for education,” in Paul Glewwe, ed., *Education Policy in Developing Countries*, Chicago: University of Chicago Press, 2013, pp. 285–338.
- Dhar, Diva, Tarun Jain, and Seema Jayachandran**, “Reshaping Adolescents’ Gender Attitudes: Evidence from a School-Based Experiment in India,” 2018.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer**, “Education, HIV, and early fertility: Experimental evidence from Kenya,” *American Economic Review*, 2015, 105 (9), 2757–97.
- Fizbein, Ariel, Norbert Schady, Francisco H.G. Ferreira, Margaret Grosh, Niall Keleher, Pedro Olinto, and Emmanuel Skoufias**, *Conditional Cash Transfers: Reducing Present and Future Poverty*, World Bank, 2009.
- Heckman, J., J. Stixrud, and S. Urzua**, “The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior,” *Journal of Labor Economics*, 2006, 24 (3), 411–482.

- Holmlund, Helena and Olmo Silva**, “Targeting noncognitive skills to improve cognitive outcomes: evidence from a remedial education intervention,” *Journal of Human Capital*, 2014, 8 (2), 126–160.
- Jensen, Robert**, “The (Perceived) Returns to Education and the Demand for Schooling,” *The Quarterly Journal of Economics*, 2010, 125 (2), 515–548.
- Lavecchia, Adam M, Heidi Liu, and Philip Oreopoulos**, “Behavioral economics of education: Progress and possibilities,” in “Handbook of the Economics of Education,” Vol. 5, Elsevier, 2016, pp. 1–74.
- Levitt, Steven D, John A List, Susanne Neckermann, and Sally Sadoff**, “The behavioralist goes to school: Leveraging behavioral economics to improve educational performance,” *American Economic Journal: Economic Policy*, 2016, 8 (4), 183–219.
- MacLeod, W Bentley and Miguel Urquiola**, “Is Education Consumption or Investment? Implications for the Effect of School Competition,” 2018.
- Muralidharan, Karthik and Nishith Prakash**, “Cycling to school: Increasing secondary school enrollment for girls in India,” *American Economic Journal: Applied Economics*, 2017, 9 (3), 321–350.
- Nguyen, Trang**, “Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar,” 2008.

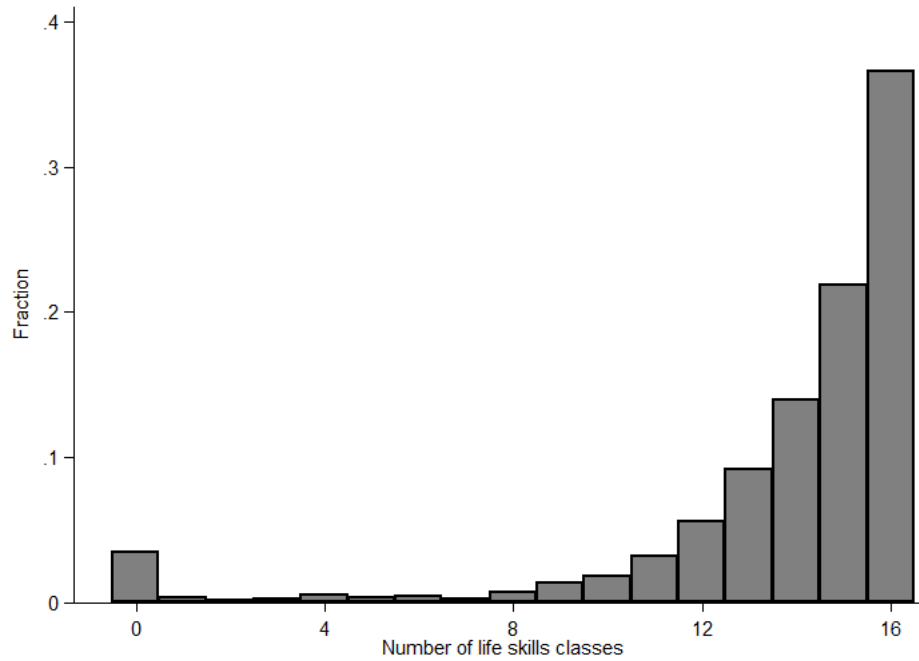


Figure 1: Number of Life Skills Classes Attended by Treatment Group Subjects in Grade 6 (out of 16 Classes)

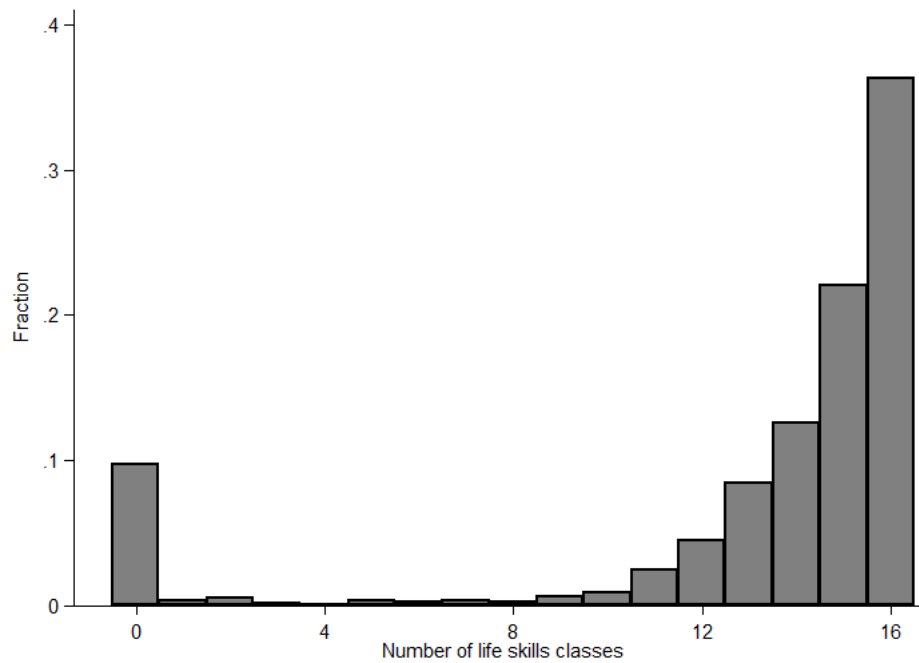


Figure 2: Number of Life Skills Classes Attended by Treatment Group Subjects in Grade 7 (out of 16 Classes)

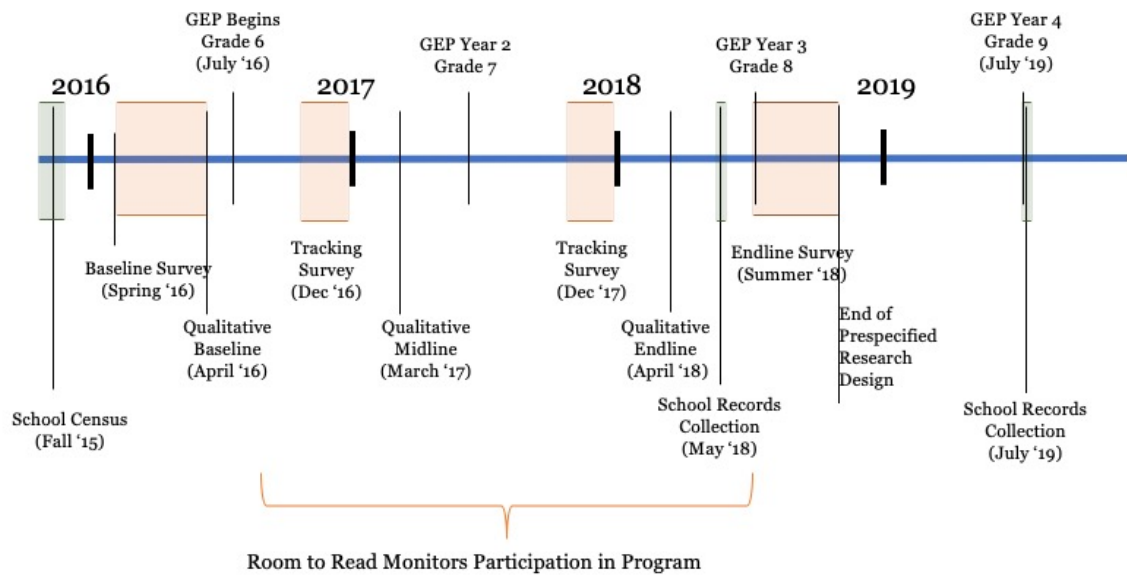


Figure 3: Intervention and Data Collection Timeline

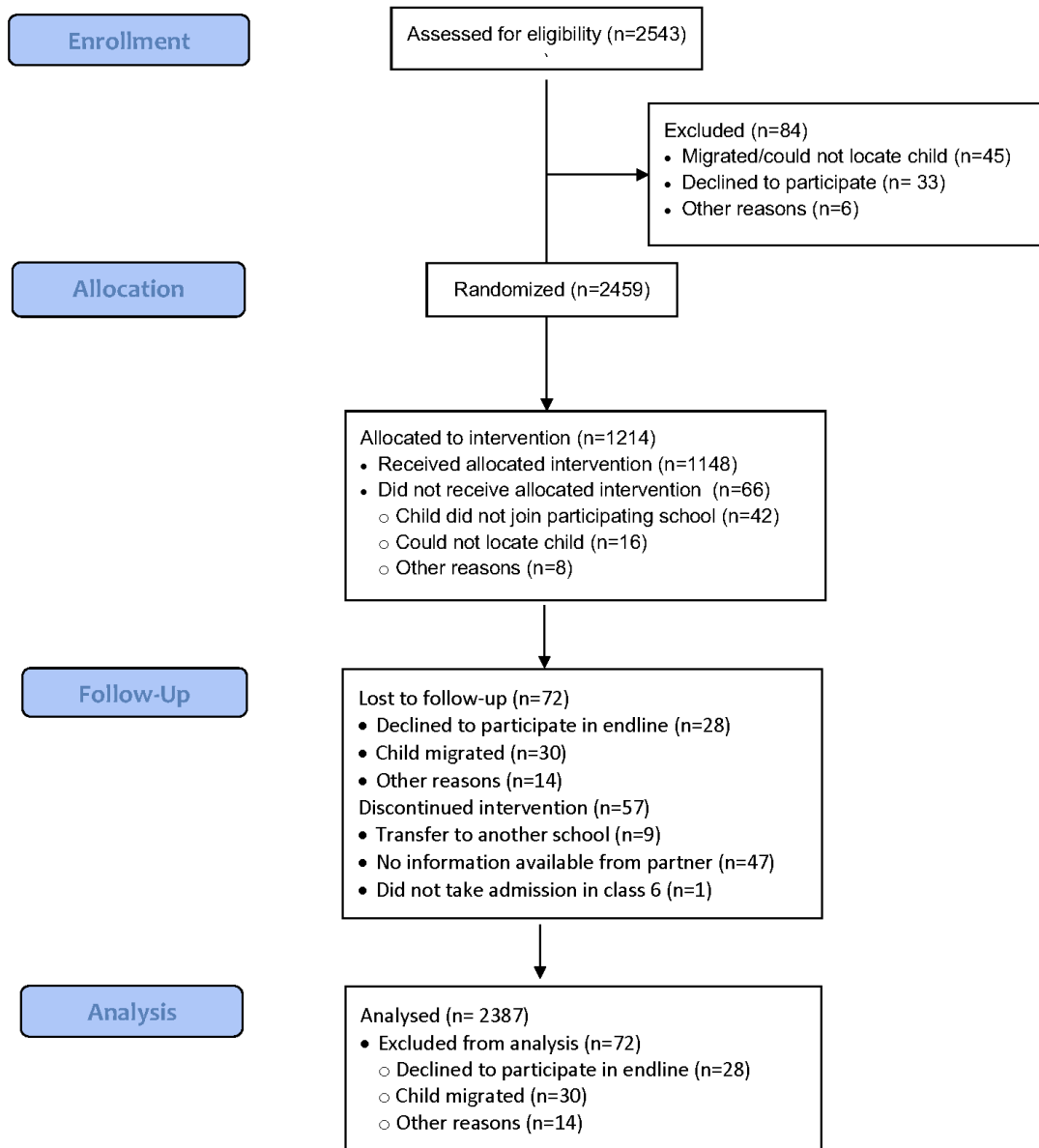


Figure 4: Flow Chart of Participants

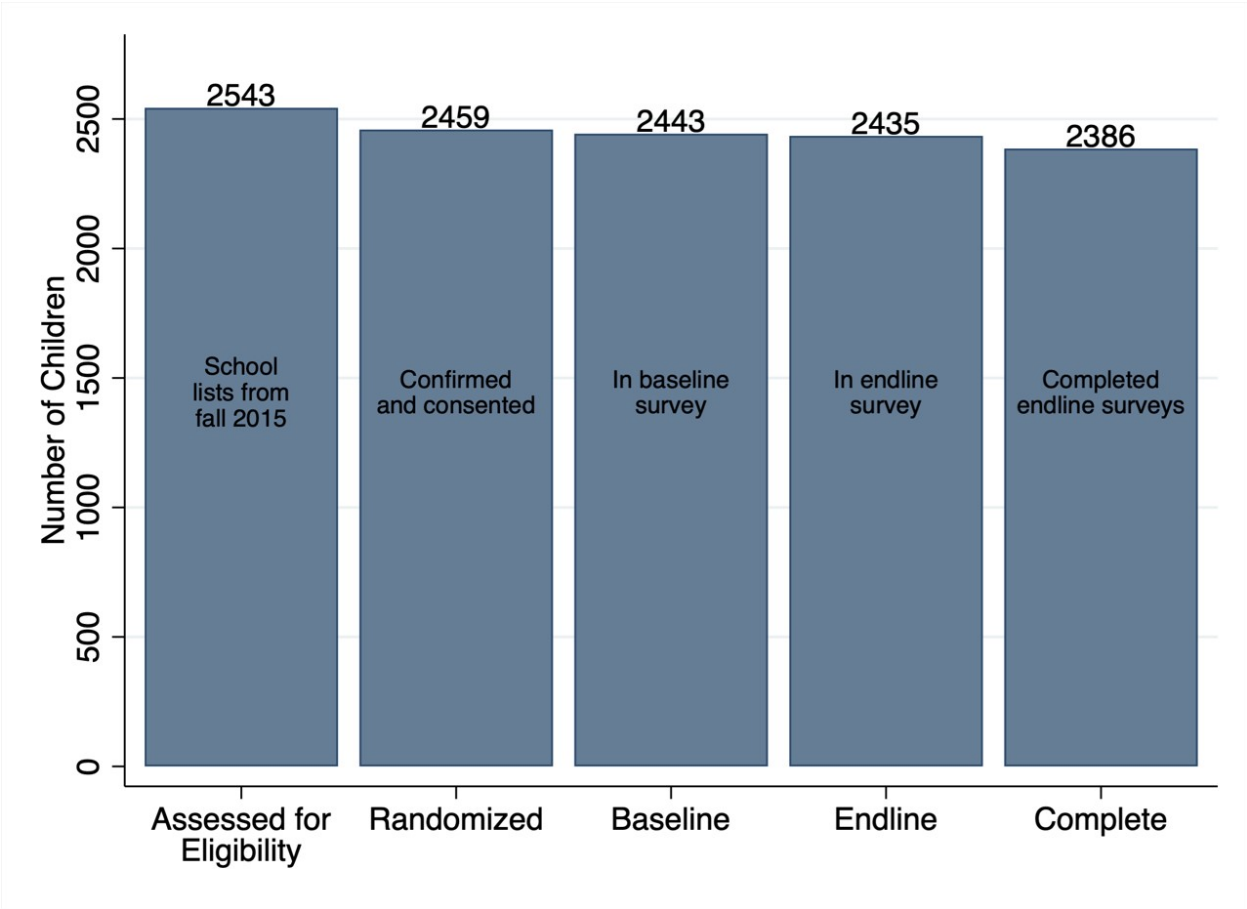


Figure 5: Attrition by Data Collection Round and Survey Type

Notes: Completed endline surveys refers to the completion of both the child and household surveys.

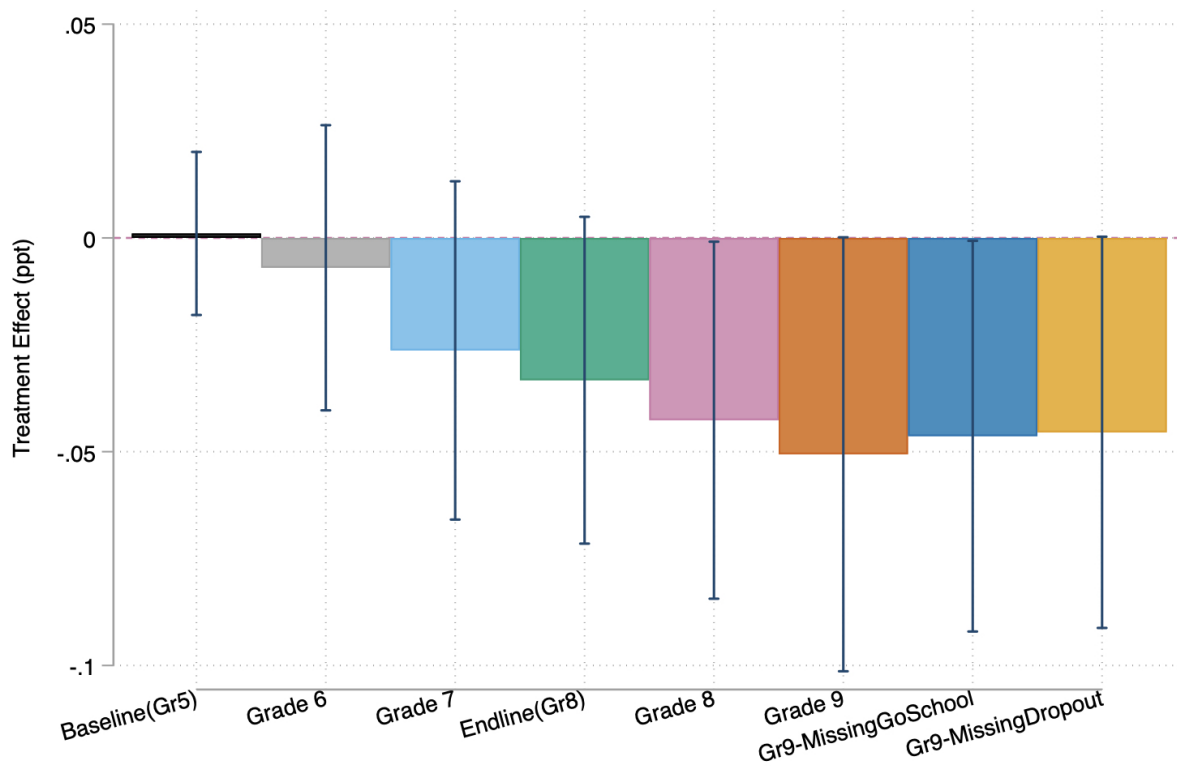


Figure 6: Treatment Effects on Dropout by Grade

Notes: This figure reports the estimated treatment effects on dropout, controlling for stratum. Baseline and endline data are from the respective surveys. All other specifications are estimated using school administrative records. Grade nine data was collected at the start of the school year and is only based on whether the child attended school in the seven days before surveying; thus all grade nine results also include controls for the number of days the school was open in the seven days prior to survey. The last two columns create bounds for the grade nine estimate by assuming all missing children went to school or did not go to school, respectively. 95 percent confidence intervals are pictured. Standard errors are clustered by school.

Table 1: Summary Statistics for Sampled Households

	Sample Mean (1)	Rajasthan Mean (2)	India Mean (3)
Number of household members	6.838	5.091	4.692
Number of boys in household (under 18)	1.379	1.005	0.834
Number of girls in household (under 18)	2.438	0.898	0.775
Enrollment: girls 10-11	97.5	92.8	95.5
Enrollment: boys 10-11	97.7	95.9	95.6
Enrollment: girls 12-14	92.0	84.0	90.2
Enrollment: boys 12-14	92.4	92.8	91.4
Marriage rate: girls 13-14	0.1095	0.0166	0.0162
Muslim	0.214	0.080	0.125
Other Backward Class	0.674	0.459	0.442
Scheduled Caste/Scheduled Tribe	0.250	0.337	0.312
Land owned (bighas)	6.283	23.924	10.664

Notes: Column (1) presents mean values averaged over all households in the study sample. Households with multiple study subjects occur as multiple observations. 16 study subjects completed a baseline child survey but no baseline household survey and thus are not represented in these summary statistics. Columns (2) and (3) present household-level mean values for respondents to the 2015-2016 Indian Demographic and Health Survey. Enrollment measures take on values from 0 to 100. Marriage rate takes on values from 0 to 1, and Muslim, Other Backward Class, and Scheduled Caste/Scheduled Tribe are all indicator variables.

Table 2: School Progression and Completion

	Survey data				Administrative data			
	Whether child has dropped out (1)	Whether child progressed to 7th grade (2)	Attendance rate (3)	Attendance dummy (4)	Dropout Grade 6 (5)	Dropout Grade 7 (6)	Dropout Grade 8 (7)	Dropout Grade 9 (8)
Panel A: Stratification controls								
Treatment	-.033* (.020)	.037* (.020)	.006 (.010)	.003 (.005)	-.007 (.017)	-.025 (.020)	-.043** (.021)	-.051* (.026)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228
R ²	.003	.004	.002	.002	.007	.011	.005	.399
Panel B: A+ Age, Economic Status, and Baseline Values								
Treatment	-.035* (.018)	.038** (.018)	.004 (.009)	.003 (.006)	-.009 (.016)	-.025 (.019)	-.044** (.020)	-.053** (.024)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228
R ²	.129	.128	.022	.014	.095	.101	.096	.433
Panel C: B+ Imbalance Variables								
Treatment	-.041** (.018)	.042** (.019)	.002 (.009)	.003 (.006)	-.008 (.016)	-.030 (.019)	-.051*** (.020)	-.056** (.024)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228
R ²	.152	.144	.027	.019	.111	.122	.114	.438
Mean Control Group	0.132	0.865	0.918	0.982	0.075	0.142	0.192	0.290

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

Column (1) uses child and household endline survey data. These data were collected at the start of eighth grade for girls who progressed one grade level each year. Columns (2) through (4) use child endline survey only. Columns (3) and (4) are conditional on school being open and child not having dropped out of school. Attendance rate in Column (3) is the fraction of school days attended in the week prior to being surveyed and the Attendance dummy in Column (4) is an indicator for having attended any days in the past week. Columns (5) through (8) rely on administrative data. In Columns (5) through (7), dropout is measured based on whether a child attended school at the conclusion of the referenced school year. In Column (8), dropout is measured based on whether a child attended school during the past week (conditional on the school being open). Column (8) includes a set of fixed effects for the number of days that the school was open in the week before administrative data collection.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table 3: Non-cognitive Skills - Survey Measures

	Socio-emotional index (1)	Freedom of movement index (2)	Empowerment index (3)	Self-esteem index (4)	Future planning index (5)	Marital expectations index (6)	Educ. / emp. aspirations index (7)	Gender norms index (8)	Cantril's ladder (9)	Enumerator assessment index (10)
Panel A: Stratification controls										
Treatment	.070*** (.023)	.020 (.022)	.094*** (.027)	.041* (.024)	.070** (.030)	-.315** (.123)	-.011 (.054)	.089*** (.034)	-.026 (.133)	.073 (.050)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380
R ²	.007	.001	.012	.003	.005	.011	.003	.009	.0006	.002
Q-statistic	0.041	0.827	0.017	0.311	0.100	0.073	0.956	0.073	0.956	0.415
Panel B: A+ Age, Economic Status, and Baseline Values										
Treatment	.063*** (.023)	.022 (.023)	.097*** (.027)	.037 (.023)	.065** (.031)	-.198** (.081)	-.0002 (.039)	.088*** (.034)	.0007 (.131)	.092* (.047)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380
R ²	.032	.015	.027	.055	.029	.304	.197	.029	.016	.05
Q-statistic	0.080	0.790	0.011	0.358	0.154	0.099	0.997	0.080	0.997	0.193
Panel C: B+ Imbalance Variables										
Treatment	.062*** (.023)	.021 (.023)	.102*** (.028)	.037 (.023)	.072** (.030)	-.174** (.082)	.014 (.047)	.093*** (.033)	.028 (.131)	.100** (.047)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380
R ²	.034	.018	.032	.059	.034	.308	.206	.036	.02	.052
Q-statistic	0.068	0.737	0.009	0.372	0.108	0.128	0.854	0.068	0.866	0.128
Mean Control Group	0.000	0.000	-0.002	-0.001	-0.016	-0.606	0.000	0.000	4.513	0.000

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent (i.e. higher values of empowerment index components all correspond to higher levels of empowerment). Marital expectations index is not mean 0 because married girls are assigned the minimum value calculated for non-married girls. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables 3 through 5 within a panel.

Table 4: Non-cognitive Skills - Demonstration Tasks and Endline Psycho-Social Indices

	Delay discounting (1)	Completed mirror drawings (2)	Mirror drawings (seconds) (3)	Scavenger hunt index (4)	Locus of control index (5)	Perceived stress index (6)	Rosenberg self-esteem index (7)
Panel A: Stratification controls							
Treatment	-0.004 (.032)	.056 (.085)	2.172 (4.472)	-.079 (.057)	-.015 (.046)	-.025 (.047)	.016 (.030)
Obs.	2380	2387	2317	2380	2380	2380	2380
R^2	.005	.003	.001	.004	.001	.001	.006
Panel B: A+ Age, Economic Status, and Baseline Values							
Treatment	-.003 (.032)	.070 (.085)	2.610 (4.535)	-.072 (.055)	-.020 (.046)	-.024 (.047)	.024 (.030)
Obs.	2380	2387	2317	2380	2380	2380	2380
R^2	.016	.02	.014	.06	.026	.011	.036
Panel C: B+ Imbalance Variables							
Treatment	.003 (.032)	.072 (.082)	2.720 (4.559)	-.065 (.055)	-.026 (.046)	-.025 (.046)	.021 (.030)
Obs.	2380	2387	2317	2380	2380	2380	2380
R^2	.021	.027	.022	.064	.028	.013	.042
Mean Control Group	0.331	3.269	119.5	0.000	0.000	0.000	0.000

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. For Columns (5) through (7) reporting measures added at endline, we control in Panels B and C for lagged values of overall life skills indices.

Delay discounting is an indicator for whether the respondent would prefer 60 Rs. in one week over 30 Rs. now (respondents were informed that they would have a chance to receive a gift valued correspondingly). Completed mirror drawings takes on values from zero to four and Mirror drawings (seconds) measures the total number of seconds spent on mirror drawings, conditional on having attempted at least one mirror drawing.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables 3 through 5 within a panel.

Table 5: Non-cognitive Skills: Parental Reports

	Parental perception of girl's strengths (1)	Parental perception of girl's self-efficacy (2)	Parental perception of freedom of movement (3)	Parent daughter communication (4)	Parental gender attitudes (5)	Parental schooling attitudes (6)	Parental marriage attitudes (7)
Panel A: Stratification controls							
Treatment	-0.042** (.018)	.004 (.029)	.021 (.029)	-.014 (.029)	.0004 (.026)	.032 (.042)	.022 (.031)
Obs.	2434	2430	2434	2434	2434	2434	2434
R ²	.004	.0001	.003	.002	.011	.003	.003
Panel B: A+ Age, Economic Status, and Baseline Values							
Treatment	-.043** (.018)	-.0007 (.030)	.025 (.028)	-.009 (.028)	.003 (.026)	.027 (.038)	.023 (.031)
Obs.	2434	2430	2434	2434	2434	2434	2434
R ²	.019	.022	.015	.025	.037	.113	.033
Panel C: B+ Imbalance Variables							
Treatment	-.040** (.018)	.010 (.029)	.031 (.028)	-.009 (.029)	.010 (.026)	.043 (.037)	.022 (.031)
Obs.	2434	2430	2434	2434	2434	2434	2434
R ²	.021	.03	.017	.027	.045	.127	.037
Mean Control Group	0.000	-0.002	0.000	0.000	0.000	0.001	-0.004

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables 3 through 5 within a panel.

Table 6: Child Labor

	Married (1)	Child works (Economically active) (2)	Child works for pay (3)	Child works outside of family (4)	Child labor (5)	Hazardous child labor (6)	Other worst forms of child labor (7)	Hours worked in a day (8)	Hours worked unpaid work (9)	Hours active (Paid + unpaid) (10)	Hour active outside house (11)
Panel A: Stratification controls											
Treatment	.042 (.029)	.049 (.040)	.021 (.025)	-.011 (.030)	.004 (.037)	.009 (.036)	.021 (.021)	.060 (.138)	.026 (.074)	.086 (.171)	.00004 (.086)
Obs. R^2	2435 .005	2386 .005	2386 .009	2387 .005	2386 .004	2386 .003	2387 .004	2386 .005	2386 .003	2386 .007	2386 .004
Panel B: A+ Age, Economic Status, and Baseline Values											
Treatment	.011 (.018)	.044 (.037)	.023 (.025)	-.008 (.029)	.006 (.034)	.012 (.033)	.021 (.020)	.001 (.120)	.006 (.069)	-.005 (.148)	-.023 (.080)
Obs. R^2	2435 .332	2386 .045	2386 .022	2387 .015	2386 .035	2386 .037	2387 .016	2386 .129	2386 .092	2386 .173	2386 .067
Panel C: B+ Imbalance Variables											
Treatment	.004 (.018)	.031 (.035)	.012 (.025)	-.008 (.030)	-.012 (.032)	-.006 (.031)	.012 (.020)	-.024 (.124)	-.009 (.069)	-.036 (.152)	-.038 (.083)
Obs. R^2	2435 .338	2386 .08	2386 .035	2387 .018	2386 .065	2386 .064	2387 .025	2386 .132	2386 .097	2386 .175	2386 .069
Mean Control Group	0.191	0.651	0.228	0.186	0.583	0.458	0.180	1.157	1.642	2.800	0.602

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

Married is an indicator variable for whether girl is married or committed (engaged). The set of survey questions used to construct each of the indicator variable outcomes in Columns (2) through (7) can be found in the analysis plan posted on-line. Time use outcomes in Columns (8) through (11) are defined based on time use patterns recorded for “a typical day in the past week.”

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table 7: Cognitive Skills

	Survey data				Administrative data			
	Hours studying at home (1)	Total hours spent at school (2)	ASER score Mathematics (3)	ASER score Hindi (4)	ASER score English (5)	GPA Grade 6 (6)	GPA Grade 7 (7)	GPA Grade 8 (8)
	Panel A: Stratification controls							
Treatment	-062 (.077)	.183 (.189)	-.021 (.077)	.032 (.093)	-.074 (.090)	-.119 (.074)	-.121 (.092)	-.033 (.083)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912
R ²	.004	.003	.004	.004	.002	.013	.006	.004
	Panel B: A+ Age, Economic Status, and Baseline Values							
Treatment	-.067 (.076)	.132 (.187)	-.032 (.070)	.008 (.089)	-.089 (.084)	-.159** (.073)	-.145 (.095)	-.028 (.086)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912
R ²	.044	.092	.073	.083	.091	.33	.215	.206
	Panel C: B+ Imbalance Variables							
Treatment	-.043 (.075)	.164 (.188)	-.014 (.070)	.022 (.089)	-.068 (.084)	-.150** (.074)	-.145 (.096)	-.026 (.087)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912
R ²	.052	.1	.078	.085	.096	.334	.216	.207
Mean Control Group	1.541	7.166	2.353	3.025	2.369	2.259	2.404	2.890

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. In Columns 3-5, controls for baseline outcome values cannot be included since cognitive tests were not conducted at baseline; Panels B and C instead include controls for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school, and grades as reported in grade five.

Time use outcomes in Columns (1) and (2) are defined based on time use patterns recorded for “a typical day in the past week.” ASER test score outcomes in Columns (3) through (5) and GPA outcomes in Columns (6) through (8) take on values between zero and four.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table 8: Understanding Channels: Child Agency and Parental Perceptions

	(1)	(2)	(3)	(4)	(5)
Panel A: Child agency					
	Sole decision-maker: Attend school	Sole decision-maker: Continue schooling	Talks to parents about marriage	Sole decision-maker: Choice of work	Index
Treatment	.068*** (.024)	.054** (.022)	.080*** (.023)	.099*** (.025)	.160*** (.035)
Obs.	2380	2380	1976	2380	2380
Mean Control group	0.412	0.328	0.270	0.375	-0.003
Panel B: Parental perceptions					
	Willing to help	Considerate	Honors adult requests	Index	
Treatment	-.040** (.020)	-.053** (.021)	-.033 (.020)	-.099*** (.037)	
Obs.	2434	2434	2434	2434	
Mean Control group	0.813	0.692	0.797	0.000	

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel A uses child survey data and Panel B uses household survey data.

Sole decision-maker: Attend school and Sole decision-maker: Continue schooling are indicators for whether the girl responds “I do/I will” when asked who mostly makes decisions about whether or not the girl will go to school and whether or not the girl will continue in school past eighth grade, respectively. Talks to parents about marriage is an indicator for whether the girl responds that she can talk to her parents about her preferences regarding who she will marry. This measure is missing for girls who are already married. Sole decision-maker: Choice of work is an indicator for whether the girl responds “I do/I will” when asked who mostly makes decisions about what type of work she will do after she finishes her studies. Willing to help is an indicator for whether the caregiver responds that it is “Certainly true” that the girl often offers to help others. Considerate is an indicator for whether the caregiver responds that it is “Certainly true” that the girl is considerate of other people’s feelings. Honors adult requests is an indicator for whether the caregiver responds that it is “Certainly true” that the girl is generally well-behaved and usually does what parents ask.

In each panel, the index is constructed from the measures in the preceding columns. To construct this index, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table 9: Understanding Channels: Social Engagement, Time Allocation, and Expectations

	(1)	(2)	(3)	(4)	(5)
Panel A: Social engagement					
	Prefers to be alone	Meets friends outside school	Has place to meet friends	Has place to stay if needed	Index
Treatment	-.053** (.023)	.042* (.023)	.065** (.028)	.034* (.017)	.100*** (.028)
Obs.	2434	2380	2380	2380	2435
Mean Control group	0.391	0.635	0.467	0.759	-0.003
Panel B: Time allocation					
	Time traveling to school	Time on mobile	Reports time on mobile	Total social time	Index
Treatment	9.066** (3.675)	.346 (.217)	.006** (.003)	9.956*** (3.659)	0.166*** (0.0585)
Obs.	2387	2387	2387	2387	2387
Mean Control group	37.15	0.092	0.0025	39.08	0.000
Panel C: Expectations					
	Wants educated job	Wants work for pay	Wants to complete secondary	Girls shouldn't complete education	Index
Treatment	.032 (.023)	.013 (.023)	.0008 (.013)	-.045** (.022)	.050 (.038)
Obs.	2387	2380	2380	2380	2387
Mean Control group	0.698	0.759	0.932	0.219	-0.002

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Column (1) of Panel A uses household survey data, and Columns (2) through (4) of Panel A as well as specifications in Panels B and C use child survey data.

Prefers to be alone is an indicator for whether the caregiver responds that it is “Certainly true” that the girl would rather be alone than with other youth. Meets friends outside school is an indicator for whether the girl responds that she has met with her friends outside of school in the last week. Has place to meet friends is an indicator for whether the girl responds that she has a place to meet her female friends at least once a week. Has place to stay if needed is an indicator for whether the girl responds that she has someone in the community who would take her in for the night if her parents were out of town and she needed a place to stay. Time allocation measures in Columns (1), (2), and (4) of Panel B are constructed based on girls’ responses regarding time spent in minutes on particular activities during a typical day in the last week. In Column (3) of Panel B, Reports time on mobile is an indicator for whether the girl reports spending any time using a mobile phone during a typical day in the last week. Wants educated job is an indicator for whether the girl responds that when she grows up she would like to work in a profession that requires completed higher secondary schooling. Wants work for pay is an indicator for whether the girl responds that she hopes to work for pay in the future. Wants to complete secondary is an indicator for whether the girl responds that she wants to complete at least secondary schooling. Girls shouldn’t complete education is an indicator for whether the girl agrees with the statement that “Since girls have to get married, they should not be sent for higher education.”

In each panel, the index is constructed from the measures in the preceding columns. To construct this index, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

A Appendix

A.1 Data Collection and Validation

Consent Process Prior to the start of each survey round, a training process focused on developing enumerator skills was undertaken. Key points included strategies to locate respondents within the community; the importance of informed consent and how to correctly structure the consent process; establishing a rapport with respondents as well as with other stakeholders in the community; maintaining fidelity to the questionnaire; full comprehension of the questionnaires themselves; and correct use of the tablets. (All data collection was implemented using ODK software on handheld tablets.)

The evaluation team enrolled individual girls and households into the evaluation sample at baseline using a detailed process of consent administered for both household and child surveys. Enumerators were trained to explain the purpose of the study, the benefits of participating, the study’s duration, and the frequency of the proposed interviews. Interviews were conducted only after respondents consented to participate and all questions regarding the study were addressed. Separate consents, both verbal and written, were obtained from the members who participated in the household survey. For the child survey, parental consent from the primary caregiver was first obtained before interviewing the child. In case the primary caregiver of the child was not available, consent was obtained from the most senior member of the household. Informed verbal consent was obtained from all children participating in the study. The consent process was then repeated for each subsequent survey.

Quantitative Data Collection The survey teams deployed to the field using household rosters that were constructed based on the lists of enrolled girls obtained from sampled schools. The information provided by the schools typically included the name of the head of household and the child herself, as well as some identifying information about the location of the household. In general, however, it was also necessary for enumerators and field supervisors to work with community members to locate each household. Field supervisors and field managers would also make courtesy visits to community stakeholders (including the sarpanch or village leader, school headmaster, and teachers) when they first arrived in the community in order to introduce the team and outline the survey’s objectives.

Each survey included a minimum of two visits to the household, as the survey administered to the girl herself was divided into two parts. This choice was made in order to maximize attention and avoid fatigue; in addition, the first visit was used to introduce a scavenger hunt task to the girl, so that she could engage in the scavenger hunt prior to the second visit. However, many households required more than two visits total to complete the data collection process, particularly as the household survey included multiple modules to be answered by different individuals. (For example, introductory modules including household rosters were administered to the head of household or the individual most knowledgeable about the household. Modules collecting information about perception of the child’s life skills were administered to the individual primarily responsible for the child’s care.)

Data Validation To minimize surveyor error, all survey skip patterns and valid response ranges were pre-programmed onto tablets prior to the start of survey activities. In addition, the survey was designed so that surveyors were required to verify that respondent identifiers and names matched our master file records prior to commencing each round of data collection. To assess data quality in real time, the project research associate was tasked with downloading collected data at

the end of each day and running a series of data quality checks in Stata to identify any survey questions generating unexpected response patterns or high rates of missing values. In addition, these data checks identified whether any surveyors were recording missing or “Don’t Know” responses with high frequency. When such cases were identified, the field staff worked with the responsible enumerator to correct surveying practices to minimize non-response.

Qualitative Data Collection Qualitative data collection was conducted at baseline, midline, and endline. This involved research activities in six schools served by Room to Read and in the associated communities. Three schools were selected in which school quality was above average, and two schools were selected in which it was below average; a sixth school was selected because it was an all girls’ school. The objective of the qualitative data collection is to understand better the channels through which the GEP changes attitudes, perceptions, and decision-making processes for girls, teachers, parents and other stakeholders. Qualitative data was collected by staff members trained in in-depth interview techniques, and collection included the transcription, translation, and coding of the resulting data.

A.2 Selection into Administrative data: Dropout and Grades

In addition to results estimated using survey data, we also present results estimated using administrative data reported on dropout and grades in Tables 2 and 7. In Appendix Tables A6, A7, and A8, we present additional robustness checks analyzing potential bias induced by selection into these administrative data.

In the analysis of school-reported data on dropout, girls are missing if the schools report no data on the girls’ whereabouts: i.e., if the girl is no longer enrolled and the school cannot identify whether she has transferred to another school (a process that requires a certificate from the originating school) or definitively dropped out. Attrition from these data is relatively infrequent in grades six through eight, but increases to 11 percent in grade nine as students are more likely to change schools prior to entering high school.

To examine the potential influence of attrition, we re-estimate the specification of interest for each grade first assuming that all missing children are not in school, and subsequently assuming that all missing children are in school. In Table A6, Column (1) reports the effect of treatment on baseline dropout (grade five), confirming there is no baseline imbalance; Columns (2) through (9) report the robustness checks for dropout in grades six, seven, eight and nine. While there is some change in estimated treatment effects, in both bounding exercises the estimated treatment effects for grade nine are not statistically distinguishable from the treatment effects for grade eight. Hence, this evidence suggests that impact of the GEP on dropout continues into high school, although we cannot say whether that effect would have persisted without the addition of material support.

For the analysis of school-reported data on test scores, scores are missing for girls who have dropped out of school as well as for other children whose missing exam scores have no singular explanation. (This is an advantage of the in-home ASER tests also conducted; missing data for the ASER scores is minimal, and restricted to those girls who were not observed in the endline survey.)

In order to analyze the potential impact of missing test scores on our findings, we first assign all missing children high and low test scores. Specifically, Column (1) of Table A7 reports the effect of treatment on baseline GPA in order to assess any baseline imbalance. In Columns (2) through (7) of the same table, we re-estimate the primary specification (1) assigning all missing children the 75th or 25th percentile GPA for children in their school. While these different assumptions about the selection into test scores move our estimates of treatment effects, the resulting treatment effect

estimates are still consistent with our hypothesis that there is no effect of treatment on in-school test scores.

In Panel A of Table A8, we examine the relationship between indicators for available test score data and treatment status. We find that treated students are less likely to have missing administrative test score data. In Panels B and C, we assess the degree to which this selection into test data would be expected to bias estimated treatment effects for administrative test score outcomes by interacting treatment status with the baseline (grade 5) administrative test score in Panel B and with the baseline attendance rate in Panel C. Interaction terms are statistically insignificant at conventional levels in all but one specification and are inconsistent in sign, suggesting that differences in missing rates as a function of treatment status are not likely to bias estimates in practice.

A.3 Bounding

Given evidence from Section 3.1.4 that girls in the control group were more likely to attrit from the endline girl survey (though not the endline household survey), we assess the potential importance of missing data in Tables A9 through A13 for those outcomes in Tables 2 through 7 that are constructed using endline girl survey responses. Specifically, we conduct separate bounding exercises corresponding to positive and negative selection. For the positive selection specifications, we assign to all missing children the 75th percentile values for index- and time use-based outcomes and the maximum response value for all other outcomes (typically indicator measures). For the negative selection specifications, we assign to all missing children the 25th percentile values for index- and time use-based outcomes and the minimum response value for all other outcomes. While these different assumptions about selection into the girl endline survey do generate some variation in our estimates of treatment effects, selection-adjusted estimates are not statistically distinguishable from the original estimates, and the statistical significance of estimates (relative to a null hypothesis of zero effect) is essentially unchanged for all included outcomes.

A.4 Heterogeneous Effects

The analysis plan pre-specified an analysis of heterogeneity along a number of dimensions: school quality, baseline child age, maternal education, and exposure of the household to recent shocks (economic shocks, crime shocks, and death/illness shocks). Heterogeneous effects for the primary outcomes of interest are reported in Tables A14 through A25 in the Appendix. In general, we fail to find evidence of significant heterogeneity in the observed treatment effects.

Appendix Tables

Table A1: Summary Statistics for Sampled Households

	Mean (1)	Std. dev. (2)	Obs. (3)
Number of sampled girls in household	1.062	0.246	2427
Number of household members	6.838	2.811	2427
Number of boys in household (under 18)	1.379	1.026	2427
Number of girls in household (under 18)	2.438	1.359	2427
Other Backward Class household	0.674	0.469	2427
Primary household source of employment = wage / salary earning	0.532	0.499	2427
Primary household source of employment = Self-employment agriculture	0.215	0.411	2427
Primary household source of employment = Self-employment non-agriculture	0.080	0.272	2427
Primary household source of employment = Casual labor in agriculture	0.013	0.114	2427
Primary household source of employment = Casual labor in non-agriculture	0.157	0.363	2427
Non-food expenditures in Rupees (last 30 days)	9906.753	4.0e+04	2427
Food expenditures in Rupees (last 30 days)	1.6e+04	2.0e+05	2427
Durables expenditures in Rupees (last year)	1.2e+05	9.6e+05	2427
Land owned (bighas)	6.283	15.959	1930
Land cultivated (bighas)	2.301	12.525	1633
Household holds NREGA card	0.756	0.430	2427
Economic shock	0.606	0.489	2427
Crime shock	0.132	0.338	2427
Death / illness shock	0.406	0.491	2427

Notes: Households with multiple study subjects occur as multiple observations. 16 study subjects completed a baseline child survey but no baseline household survey and thus are not represented in these baseline summary statistics.

Primary household source of employment measures are indicator variables. 7% of households, or 182 households, report that they own no land individually but access collectively owned land. 315 households, or 13%, cannot estimate the amount of land owned. 8% of households, or 206 households, do not report land cultivated because it is cultivated collectively, and an additional 588 households (or 24%) cannot estimate the amount of land cultivated. Economic shock is an indicator for loss of employment or lowered income of any household member or bankruptcy of family business in last 12 months. Crime shock is an indicator for having experienced robbery, assault, physical aggression, a land dispute, or a family dispute in last 12 months. Death/illness shock is an indicator for death, serious illness, or accident of a household member in last 12 months.

Table A2: Summary Statistics for Sampled Children

	Mean (1)	Std. dev. (2)	Obs. (3)
Stratification (Baseline school characteristics):			
Below median school quality	0.509	0.500	2407
Above median school quality	0.491	0.500	2407
Subject characteristics:			
Age	10.989	1.425	2419
Maternal education (1=completed primary or above)	0.172	0.377	2426
Girl's marital status (1=married)	0.167	0.373	2421
Child has dropped out of school	0.025	0.156	2440
Child is in grade five	0.975	0.156	2399
Any attendance in last week (conditional on not dropping out)	0.889	0.314	2026
Attendance rate in last week (conditional on attendance)	0.788	0.337	2026
Delay discounting	0.178	0.383	2399
Completed mirror drawings	2.396	1.248	2415
Mirror drawings (seconds)	68.407	70.266	2399
Scavenger hunt index	-0.023	0.969	2398
Socio-emotional index	0.017	0.464	2399
Freedom of movement index	-0.001	0.602	2399
Empowerment index	-0.004	0.416	2399
Self-esteem index	0.013	0.488	2399
Future planning index	0.051	0.601	2399
Marital expectations index	-0.496	1.435	2399
Education / employment aspirations index	-0.017	0.794	2399
Gender norms index	-0.003	0.509	2399
Cantril's ladder	7.955	2.418	2399
Enumerator assessment index	-0.027	0.891	2399
Parental perception of girl's strengths	0.004	0.365	2425
Parental perception freedom of movement	-0.021	0.59	2425
Parent-daughter communication	0.002	0.422	2443
Parental gender attitudes	0.001	0.432	2425
Parental schooling attitudes	0.007	0.695	2427
Parental marriage attitudes	-0.005	0.516	2425
Child works	0.914	0.28	2398
Child works for pay	0.844	0.363	2398
Child works outside of family activity	0.697	0.460	2399
Child labor	0.874	0.332	2398
Hazardous child labor	0.642	0.479	2397
Other worst forms of child labor	0.225	0.418	2399
Hours economically active in a day	1.052	1.691	2397
Hours in unpaid household services in a day	1.447	1.448	2397
Total hours active	2.499	2.308	2397
Hours active outside house	0.825	1.454	2397
Hours studying at home	0.704	0.955	2397
Total hours spent on school	6.105	2.823	2397

Notes: One household did not complete a roster and thus is not represented in these baseline summary statistics.

Maternal education is measured at endline and is missing if child is not present in endline survey. Any attendance in last week is missing if child has dropped out or her school was not open in past week. Attendance rate in last week is missing if child has dropped out, her school was not open in past week, or she did not attend school in past week. Details regarding the remaining variables and indices can be found in the analysis plan posted on-line.

Table A3: Balance Tests for Household Variables

	Control		Treatment		Difference		Q-stat (7)
	Mean (1)	Std. dev. (2)	Mean (3)	Std. dev. (4)	Coef. (5)	Std. error (6)	
Number of sampled girls in household	1.057	0.241	1.067	0.251	0.010	(0.015)	0.834
Number of household members	6.893	2.763	6.781	2.860	-0.106	(0.140)	0.829
Number of boys in household (under 18)	1.358	1.047	1.402	1.003	0.046	(0.050)	0.745
Number of girls in household (under 18)	2.456	1.340	2.419	1.380	-0.037	(0.063)	0.834
Other backward castes household	0.631	0.483	0.720	0.449	0.088**	(0.038)	0.579
Primary household source of employment = wage / salary earning	0.536	0.499	0.527	0.499	-0.012	(0.031)	0.910
Primary household source of employment = Self-employment agriculture	0.210	0.407	0.220	0.415	0.011	(0.032)	0.930
Primary household source of employment = Self-employment non-agriculture	0.072	0.258	0.089	0.285	0.018	(0.015)	0.655
Primary household source of employment = Casual labor in agriculture	0.015	0.120	0.012	0.108	-0.003	(0.005)	0.834
Primary household source of employment = Casual labor in non-agriculture	0.162	0.369	0.151	0.358	-0.010	(0.018)	0.834
Non-food expenditures in Rupees (last 30 days)	1.0e+04	5.4e+04	9453.617	1.8e+04	-879.716	(1678.969)	0.834
Food expenditures in Rupees (last 30 days)	2.2e+04	2.9e+05	1.0e+04	1.1e+04	-1.2e+04	(7903.035)	0.655
Durables expenditures in Rupees (last year)	1.1e+05	5.5e+05	1.4e+05	1.3e+06	3.3e+04	(4.0e+04)	0.784
Land owned (bighas)	5.653	11.828	6.901	19.153	1.246	(1.104)	0.655
Land cultivated (bighas)	2.069	8.285	2.540	15.738	0.455	(0.743)	0.834
Household holds NREGA card	0.712	0.453	0.802	0.399	0.090	(0.063)	0.655
Economic shock	0.593	0.491	0.620	0.486	0.027	(0.024)	0.655
Crime shock	0.126	0.332	0.138	0.345	0.011	(0.017)	0.834
Death / illness shock	0.396	0.489	0.417	0.493	0.023	(0.021)	0.655

Notes: Households with multiple study subjects occur as multiple observations. 16 study subjects completed a baseline child survey but no baseline household survey and thus are not represented in these baseline summary statistics.

Primary household source of employment measures are indicator variables. 7% of households, or 182 households, report that they own no land individually but access collectively owned land. 315 households, or 13%, cannot estimate the amount of land owned. 8% of households, or 206 households, do not report land cultivated because it is cultivated collectively, and an additional 588 households (or 24%) cannot estimate the amount of land cultivated. Economic shock is an indicator for loss of employment or lowered income of any household member or bankruptcy of family business in last 12 months. Crime shock is an indicator for having experienced robbery, assault, physical aggression, a land dispute, or a family dispute in last 12 months. Death/illness shock is an indicator for death, serious illness, or accident of a household member in last 12 months.

The columns under the header "Difference" report the result of the regression of the row variable on an indicator for treatment and stratification fixed effects. Standard errors are clustered by school.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables A3 through A5 within a panel.

Table A4: Balance Tests for Child Variables

	Control	Treatment	Difference	Q-stat
	Mean	Mean	Coef.	stat
	Std. dev.	Std. dev.	error	(7)
	(1)	(3)	(5)	(6)
Stratification (Baseline school characteristics):				
Below median school quality	0.515	0.502	0.000	(—)
Above median school quality	0.485	0.498	0.000	(—)
Subject characteristics:				
Age	10.960	11.019	0.058	(0.081)
Maternal education (1=completed primary or above)	0.187	0.156	-0.030	(0.023)
Girl's marital status (1=married)	0.141	0.194	0.053*	(0.028)
Child has dropped out of school	0.024	0.026	0.001	(0.010)
Child is in grade five	0.975	0.974	-0.000	(0.010)
Any attendance in last week (conditional on not dropping out)	0.870	0.908	0.040	(0.025)
Attendance rate in last week (conditional on attendance)	0.768	0.808	0.044	(0.028)
Delay discounting	0.171	0.186	0.013	(0.026)
Completed mirror drawings	2.489	2.331	-0.149	(0.117)
Mirror drawings (seconds)	69.452	67.318	-1.833	(6.043)
Scavenger hunt index	-0.000	-0.048	-0.045	(0.076)
Socio-emotional index	-0.000	0.035	0.034	(0.030)
Freedom of movement index	0.000	-0.002	-0.004	(0.046)
Empowerment index	-0.000	-0.008	-0.007	(0.029)
Self-esteem index	0.000	0.027	0.027	(0.030)
Future planning index	0.020	0.084	0.062*	(0.033)
Marital expectations index	-0.401	-0.595	-0.192*	(0.104)
Education / employment aspirations index	-0.002	-0.033	-0.029	(0.053)
Gender norms index	-0.000	-0.005	-0.005	(0.034)
Cantril's ladder	8.029	7.877	-0.154	(0.152)
Enumerator assessment index	-0.000	-0.056	-0.055	(0.052)

Notes: One household did not complete a roster and thus is not represented in these baseline summary statistics.

Maternal education is measured at endline and is missing if child is not present in endline survey. Any attendance in last week is missing if child has dropped out or her school was not open in past week. Attendance rate in last week is missing if child has dropped out, her school was not open in past week, or she did not attend school in past week. Details regarding the remaining variables and indices can be found in the analysis plan posted on-line.

The columns under the header "Difference" report the result of the regression of the row variable on an indicator for treatment and stratification fixed effects. Standard errors are clustered by school.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables A3 through A5 within a panel.

Table A5: Balance Tests for Child Variables, cont.

	Control	Treatment	Difference	Q-stat
	Mean	Mean	Coef.	stat
	(1)	(3)	(5)	(7)
	Std. dev.	Std. dev.	Std. error	
	(2)	(4)	(6)	
Parental perception of girl's strengths	-0.000	0.007	0.008	0.955
Parental perception of girl's self-efficacy	0.000	0.048	0.047	0.655
Parental perception freedom of movement	-0.000	-0.043	-0.044	0.655
Parent-daughter communication	0.001	0.002	0.002	0.972
Parental gender attitudes	-0.000	0.003	0.003	0.967
Parental schooling attitudes	0.003	0.012	0.010	0.955
Parental marriage attitudes	-0.005	-0.005	0.000	0.999
Child works	0.884	0.945	0.060***	0.234
Child works for pay	0.829	0.859	0.029	0.655
Child works outside of family activity	0.674	0.721	0.046	0.655
Child labor	0.855	0.893	0.035	0.655
Hazardous child labor	0.620	0.665	0.042	0.655
Other worst forms of child labor	0.219	0.231	0.010	0.910
Hours economically active in a day	0.945	1.164	0.217*	0.632
Hours in unpaid household services in a day	1.415	1.480	0.065	0.721
Total hours active	2.360	2.644	0.282*	0.579
Hours active outside house	0.719	0.935	0.215**	0.579
Hours studying at home	0.713	0.694	-0.013	0.955
Total hours spent on school	6.014	6.199	0.192	0.829

Notes: One household did not complete a roster and thus is not represented in these baseline summary statistics.

Details regarding the included variables and indices can be found in the analysis plan posted on-line.

The columns under the header "Difference" report the result of the regression of the row variable on an indicator for treatment and stratification fixed effects. Standard errors are clustered by school.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Q-statistics are False Discovery Rate corrected q-values based on Benjamini and Hochberg (1995). These are computed by pooling all specifications included in Tables A3 through A5 within a panel.

Table A6: Robustness Checks for Dropout Data

	5th grade		6th grade		7th grade		8th grade		9th grade	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Treatment	.001 (.010)	-.003 (.017)	-.032 (.022)	-.020 (.019)	-.053** (.024)	-.032* (.019)	-.063** (.025)	-.046** (.023)	-.045* (.023)	
Obs.	2459	2459	2459	2459	2459	2459	2459	2459	2459	
R ²	.005	.007	.005	.009	.006	.004	.007	.272	.511	
Panel A: Stratification controls										
Treatment	-.005 (.016)	-.033 (.022)	-.023 (.018)	-.053** (.023)	-.033* (.018)	-.064*** (.024)	-.049** (.022)	-.048** (.022)		
Obs.	2459	2459	2459	2459	2459	2459	2459	2459		
R ²	.066	.056	.083	.065	.091	.079	.309	.537		
Panel B: A+ Age, Economic Status, and Baseline Values										
Treatment	-.004 (.016)	-.034 (.022)	-.027 (.018)	-.056** (.023)	-.040** (.018)	-.066*** (.023)	-.052** (.022)	-.051** (.022)		
Obs.	2459	2459	2459	2459	2459	2459	2459	2459		
R ²	.078	.067	.094	.08	.109	.094	.314	.54		
Panel C: B+ Imbalance Variables										

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

All outcome measures are constructed using administrative data. Column 1 uses dropout by end of fifth grade as a reference outcome. The remaining columns create bounds to assess the importance of missing data by assuming all missing children did not attend school (in even-numbered columns) or attended school (in odd-numbered columns).

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A7: Robustness Checks for Administrative Test Data

	5th (1)	6th Missing (2)	6th Missing 25th (3)	7th Missing 75th (4)	7th Missing 25th (5)	8th Missing 75th (6)	8th Missing 25th (7)
Panel A: Stratification controls							
Treatment	.023 (.075)	-.129* (.075)	-.111 (.072)	-.132 (.091)	-.085 (.085)	-.053 (.082)	.016 (.087)
Obs.	2356	2459	2459	2459	2459	2459	2459
R ²	.019	.013	.014	.008	.003	.008	.005
Panel B: A+ Age, Economic Status, and Baseline Values							
Treatment		-.137* (.073)	-.142* (.072)	-.133 (.092)	-.109 (.088)	-.042 (.082)	.005 (.088)
Obs.		2459	2459	2459	2459	2459	2459
R ²		.266	.31	.144	.176	.124	.173
Panel C: B+ Imbalance Variables							
Treatment		-.130* (.074)	-.134* (.073)	-.131 (.093)	-.103 (.089)	-.045 (.083)	.006 (.088)
Obs.		2459	2459	2459	2459	2459	2459
R ²		.268	.313	.145	.178	.126	.177
Mean Control Group	3.064	2.312	2.200	2.500	2.303	3.007	2.757

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables.

All outcome measures are constructed using administrative data. Column 1 uses fifth grade test score data as a reference outcome. The remaining columns create bounds to assess the importance of missing data by assuming all missing children would have scored at the 75th percentile of the test score distribution (in even-numbered columns) or at the 25th percentile (in odd-numbered columns).

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A8: Selection into Survey-Based and Administrative Test Data

	ASER Score Available (1)	GPA Grade 6 Available (2)	GPA Grade 7 Available (3)	GPA Grade 8 Available (4)
Panel A: Selection into Test Data (Stratification controls)				
Treatment	0.011 (0.008)	0.032 (0.022)	0.053** (0.024)	0.069*** (0.026)
Obs.	2459	2459	2459	2459
R^2	.0018	.0051	.0064	.0085
Panel B: Selection into Test Data by Grade 5 GPA (Stratification controls)				
Treatment	0.048 (0.032)	-0.143* (0.073)	-0.087 (0.098)	0.068 (0.111)
Treatment * Grade 5 GPA	-0.011 (0.010)	0.053** (0.023)	0.042 (0.031)	-0.003 (0.033)
Grade 5 GPA	0.011 (0.008)	-0.007 (0.017)	0.036 (0.023)	0.086*** (0.025)
Obs.	2356	2356	2356	2356
R^2	.0033	.0089	.0170	.0287
Panel C: Selection into Test Data by Grade 5 Attendance (Stratification controls)				
Treatment	0.036 (0.022)	0.053 (0.056)	0.054 (0.063)	0.148** (0.070)
Treatment * Grade 5 Attendance	-0.031 (0.025)	-0.020 (0.060)	0.010 (0.072)	-0.099 (0.077)
Grade 5 Attendance	0.036* (0.019)	0.134*** (0.048)	0.168*** (0.054)	0.266*** (0.058)
Obs.	2026	2026	2026	2026
R^2	.0053	.0252	.0319	.0470
Mean Control Group	.962	.870	.778	.744

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panels B and C contain results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics (reported), and stratification fixed effects. Missing observations in Panels B and C correspond to missing baseline values of the specified characteristic. The dependent variable in Column 1 is an indicator for whether survey-administered ASER test score data is available. The dependent variables in Columns 2-4 are indicators for whether administrative test score data from grades 6-8 is available.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A9: School Progression and Completion (Attrition Bounds)

	Whether child has dropped out (1)	Whether child progressed to 7th grade (2)	Attendance rate (3)	Attendance dummy (4)
Panel A: Stratification controls				
Treatment (negative selection)	-.032* (.019)	.048** (.021)	.018 (.015)	.017 (.014)
Treatment (positive selection)	-.037* (.020)	.034* (.020)	.004 (.009)	.003 (.005)
Panel B: A+ Age, Economic Status, and Baseline Values				
Treatment (negative selection)	-.034* (.018)	.051*** (.019)	.014 (.015)	.014 (.014)
Treatment (positive selection)	-.039** (.018)	.034* (.018)	.002 (.009)	.003 (.005)
Panel C: B+ Imbalance Variables				
Treatment (negative selection)	-.041** (.018)	.055*** (.019)	.014 (.015)	.016 (.014)
Treatment (positive selection)	-.045** (.019)	.039** (.018)	.0008 (.009)	.002 (.006)
Mean Control Group (negative selection)	0.130	0.834	0.874	0.934
Mean Control Group (positive selection)	0.143	0.870	0.922	0.982

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. Column 1 uses child and household endline survey data. These data were collected at the start of eighth grade for girls who progressed one grade level each year. Columns 2-4 use child endline survey only. Columns 3 - 4 are conditional on school being open and child not having dropped out of school. Attendance rate in Column 3 is the fraction of school days attended in the week prior to being surveyed and the Attendance dummy in Column 4 is an indicator for having attended any days in the past week. Columns 1-2 include 2,459 observations and Columns 3-4 include 2,178 observations (since children who have dropped out of school are excluded).

To construct bounds based on negative selection, missing observations are set equal to zero (the minimum value) for each included outcome. To construct bounds based on positive selection, missing observations are set equal to one (the maximum value) for each included outcome.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A10: Non-cognitive Skills - Survey Measures (Attrition Bounds)

	Socio-emotional index (1)	Freedom of movement index (2)	Empowerment index (3)	Self-esteem index (4)	Future planning index (5)	Marital expectations index (6)	Educ. / emp. aspirations index (7)	Gender norms index (8)	Cantril's ladder (9)	Enumerator assessment index (10)
Panel A: Stratification controls										
Treatment (negative selection)	.071*** (.023)	.018 (.022)	.094*** (.026)	.042* (.023)	.073** (.029)	-.305** (.119)	-.008 (.052)	.089*** (.033)	-.008 (.129)	.077 (.049)
Treatment (positive selection)	.064*** (.023)	.018 (.022)	.087*** (.026)	.036 (.023)	.062** (.030)	-.316*** (.119)	-.021 (.053)	.080** (.033)	-.031 (.129)	.065 (.049)
Panel B: A+ Age, Economic Status, and Baseline Values										
Treatment (negative selection)	.064*** (.023)	.019 (.022)	.097*** (.026)	.037* (.022)	.066** (.029)	-.197** (.078)	-.00007 (.038)	.090*** (.032)	.022 (.127)	.096** (.046)
Treatment (positive selection)	.056** (.022)	.019 (.022)	.088*** (.026)	.031 (.022)	.055* (.030)	-.211*** (.079)	-.016 (.040)	.079** (.033)	-.007 (.128)	.081* (.046)
Panel C: B+ Imbalance Variables										
Treatment (negative selection)	.063*** (.022)	.018 (.022)	.102*** (.027)	.038* (.023)	.073*** (.028)	-.173** (.079)	.018 (.046)	.095*** (.032)	.047 (.127)	.104** (.046)
Treatment (positive selection)	.056** (.022)	.018 (.022)	.094*** (.027)	.032 (.023)	.063** (.029)	-.186** (.080)	.004 (.047)	.085*** (.032)	.021 (.127)	.090** (.046)
Mean Control Group (negative selection)	-0.0093	0.0048	-0.012	-0.007	-0.029	-0.611	-0.008	-0.009	4.455	-0.019
Mean Control Group (positive selection)	0.013	0.0048	0.012	0.011	0.005	-0.575	0.034	0.021	4.531	0.022

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. All columns include 2,459 observations.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent (i.e. higher values of empowerment index components all correspond to higher levels of empowerment). Marital expectations index is not mean 0 because married girls are assigned the minimum value calculated for non-married girls. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

To construct bounds based on negative selection, missing observations are set equal to the 25th percentile value for each included outcome. To construct bounds based on positive selection, missing observations are set equal to the 75th percentile value for each included outcome.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A11: Non-cognitive Skills - Demonstration Tasks and Endline Psycho-Social Indices (Attrition Bounds)

	Delay discounting (1)	Completed mirror drawings (2)	Mirror drawings (seconds) (3)	Scavenger hunt index (4)	Locus of control (5)	Perceived stress index (6)	Rosenberg self-esteem index (7)
Panel A: Stratification controls							
Treatment (negative selection)	.003 (.032)	.099 (.090)	2.172 (4.472)	-.066 (.056)	-.007 (.045)	-.016 (.046)	.019 (.028)
Treatment (positive selection)	-.008 (.030)	.044 (.082)	1.577 (4.331)	-.085 (.056)	-.023 (.045)	-.028 (.045)	.012 (.029)
Panel B: A+ Age, Economic Status, and Baseline Values							
Treatment (negative selection)	.001 (.031)	.116 (.088)	2.610 (4.535)	-.054 (.054)	-.007 (.045)	-.013 (.046)	.027 (.028)
Treatment (positive selection)	-.013 (.030)	.055 (.082)	1.775 (4.394)	-.079 (.054)	-.026 (.044)	-.027 (.045)	.019 (.029)
Panel C: B+ Imbalance Variables							
Treatment (negative selection)	.007 (.031)	.115 (.084)	2.720 (4.559)	-.048 (.053)	-.014 (.045)	-.015 (.046)	.025 (.028)
Treatment (positive selection)	-.006 (.030)	.058 (.079)	1.959 (4.433)	-.071 (.054)	-.031 (.044)	-.027 (.045)	.017 (.029)
Mean Control Group (negative selection)	0.319	3.151	119.5	-0.037	-0.029	-0.028	-0.012
Mean Control Group (positive selection)	0.357	3.300	121.0	0.029	0.026	0.011	0.011

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. For Columns 5-7 measures added at endline, we control in Panels B and C for lagged values of overall life skills indices. Columns 1, 2 and 4-7 include 2,459 observations. Column 3 include 2,317 observations for negative selection imputation and 2,389 observations for positive selection imputation (this outcome is missing for children who did not complete any mirror drawings and the number of children with any completed drawings varies based on the imputation approach).

Delay discounting is an indicator for whether the respondent would prefer 60 Rs. in one week over 30 Rs. now (respondents were informed that they would have a chance to receive a gift valued correspondingly). Completed mirror drawings takes on values from 0 to 4 and Mirror drawings (seconds) measures the total number of seconds spent on mirror drawings, conditional on having attempted at least one mirror drawing.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

To construct bounds based on negative selection, missing observations are set equal to zero (the minimum value) in Columns (1)-(2) and to the 25th percentile value in Columns (3)-(7). To construct bounds based on positive selection, missing observations are set equal to the maximum value in Columns (1)-(2) and to the 75th percentile value in Columns (3)-(7).

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A13: Cognitive Skills (Attrition Bounds)

	Hours studying at home (1)	Total hours spent at school (2)	ASER score Mathematics (3)	ASER score Hindi (4)	ASER score English (5)
Panel A: Stratification controls					
Treatment (negative selection)	-.046 (.075)	.182 (.183)	.006 (.077)	.065 (.092)	-.045 (.088)
Treatment (positive selection)	-.067 (.074)	.152 (.183)	-.039 (.075)	.020 (.091)	-.090 (.089)
Panel B: A+ Age, Economic Status, and Baseline Values					
Treatment (negative selection)	-.046 (.074)	.139 (.181)	-.004 (.069)	.044 (.086)	-.060 (.081)
Treatment (positive selection)	-.072 (.073)	.104 (.182)	-.054 (.070)	-.006 (.088)	-.110 (.084)
Panel C: B+ Imbalance Variables					
Treatment (negative selection)	-.022 (.073)	.170 (.182)	.011 (.070)	.053 (.087)	-.042 (.081)
Treatment (positive selection)	-.047 (.073)	.137 (.183)	-.033 (.069)	.009 (.088)	-.087 (.084)
Mean Control Group (negative selection)	1.503	7.157	2.264	2.911	2.280
Mean Control Group (positive selection)	1.558	7.234	2.415	3.062	2.431

Notes: Panel A contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported) and stratification fixed effects. Panel B adds age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. Panel C adds in controls for variables that appear imbalanced in the balance tables. In Columns 3-5, controls for baseline outcome values cannot be included since cognitive tests were not conducted at baseline; Panels B and C instead include controls for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school, and grades as reported in grade five. All columns include 2,459 observations.

Time use outcomes in Columns 1-2 are defined based on time use patterns recorded for “a typical day in the past week.” ASER test score outcomes in Columns 3-5 take on values between 0 and 4.

To construct bounds based on negative selection, missing observations are set equal to the 25th percentile value in Columns (1)-(2) and to the minimum value in Columns (3)-(5). To construct bounds based on positive selection, missing observations are set equal to the 75th percentile value in Columns (1)-(2) and to the maximum value in Columns (3)-(5).

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A14: School Progression and Completion: Heterogeneous Effects

	Survey data				Administrative data			
	Whether child has dropped out (1)	Whether child progressed to 7th grade (2)	Attendance rate (3)	Attendance dummy (4)	Dropout Grade 6 (5)	Dropout Grade 7 (6)	Dropout Grade 8 (7)	Dropout Grade 9 (8)
	Panel A: School quality							
Treatment	-0.16 (.026)	.020 (.027)	-0.01 (.013)	.0001 (.007)	-.025 (.030)	-.044 (.033)	-.039 (.032)	-.070 (.054)
Treatment int	-.033 (.039)	.034 (.041)	.014 (.020)	.006 (.011)	.035 (.035)	.038 (.041)	-.010 (.043)	.128 (.095)
Obs.	2397	2351	2058	2058	2338	2287	2419	2196
	Panel B: Baseline age							
Treatment	.164 (.103)	-.113 (.103)	-.060 (.041)	-.025 (.028)	.115 (.073)	.191* (.108)	.116 (.116)	.202 (.161)
Treatment int	-.018* (.010)	.014 (.010)	.006 (.004)	.003 (.003)	-.011 (.007)	-.020* (.010)	-.014 (.011)	-.019 (.015)
Obs.	2431	2385	2087	2087	2372	2317	2453	2226
	Panel C: Maternal education							
Treatment	-.034 (.023)	.037 (.024)	-.005 (.010)	-2.87e-06 (.007)	-.010 (.019)	-.030 (.023)	-.047* (.025)	.010 (.052)
Treatment int	.00006 (.036)	.008 (.040)	.057*** (.021)	.019 (.012)	.003 (.030)	.023 (.039)	.030 (.041)	-.064 (.056)
Obs.	2424	2378	2081	2081	2346	2291	2426	2200

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

Column (1) uses child and household endline survey data. These data were collected at the start of eighth grade for girls who progressed one grade level each year. Columns (2) through (4) use child endline survey only. Columns (3) and (4) are conditional on school being open and child not having dropped out of school. Attendance rate in Column (3) is the fraction of school days attended in the week prior to being surveyed and the Attendance dummy in Column (4) is an indicator for having attended any days in the past week. Columns (5) through (8) rely on administrative data. In Columns (5) through (7), dropout is measured based on whether a child attended school at the conclusion of the referenced school year. In Column (8), dropout is measured based on whether a child attended school during the past week (conditional on the school being open). Column (8) includes a set of fixed effects for the number of days that the school was open in the week before administrative data collection.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A15: School Progression and Completion: Heterogeneous Effects for Household Shocks

	Survey data				Administrative data			
	Whether child has dropped out (1)	Whether child progressed to 7th grade (2)	Attendance rate (3)	Attendance dummy (4)	Dropout Grade 6 (5)	Dropout Grade 7 (6)	Dropout Grade 8 (7)	Dropout Grade 9 (8)
	Panel A: Economic shock							
Treatment	-0.025 (.024)	.040 (.024)	.012 (.016)	-.0006 (.011)	-.0005 (.021)	-.010 (.028)	-.021 (.027)	.030 (.051)
Treatment int	-.015 (.025)	-.004 (.025)	-.011 (.019)	.006 (.013)	-.011 (.022)	-.024 (.032)	-.036 (.031)	-.054 (.043)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228
	Panel B: Crime shock							
Treatment	-.029 (.021)	.032 (.022)	.015 (.010)	.004 (.006)	-.001 (.018)	-.012 (.021)	-.034 (.022)	.003 (.048)
Treatment int	-.034 (.038)	.031 (.038)	-.071*** (.025)	-.010 (.013)	-.048 (.029)	-.099** (.042)	-.068 (.049)	-.038 (.066)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228
	Panel C: Death/illness shock							
Treatment	-.045* (.025)	.049* (.026)	.006 (.012)	.002 (.007)	-.004 (.018)	-.018 (.023)	-.046* (.026)	.036 (.053)
Treatment int	.030 (.027)	-.030 (.029)	-.001 (.017)	.004 (.010)	-.008 (.022)	-.016 (.030)	.008 (.031)	-.093** (.041)
Obs.	2433	2387	2089	2089	2374	2319	2455	2228

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

Column (1) uses child and household endline survey data. These data were collected at the start of eighth grade for girls who progressed one grade level each year. Columns (2) through (4) use child endline survey only. Columns (3) and (4) are conditional on school being open and child not having dropped out of school. Attendance rate in Column (3) is the fraction of school days attended in the week prior to being surveyed and the Attendance dummy in Column (4) is an indicator for having attended any days in the past week. Columns (5) through (8) rely on administrative data. In Columns (5) through (7), dropout is measured based on whether a child attended school at the conclusion of the referenced school year. In Column (8), dropout is measured based on whether a child attended school during the past week (conditional on the school being open). Column (8) includes a set of fixed effects for the number of days that the school was open in the week before administrative data collection.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A16: Non-cognitive Skills, Survey Measures: Heterogeneous Effects

	Socio-emotional index (1)	Freedom of movement index (2)	Empowerment index (3)	Self-esteem index (4)	Future planning index (5)	Marital expectations index (6)	Educ. / emp. aspirations index (7)	Gender norms index (8)	Cantril's ladder (9)	Enumerator assessment index (10)
Panel A: School quality										
Treatment	.062** (.028)	-.026 (.032)	.133*** (.041)	.065* (.034)	.080* (.045)	-.482** (.188)	-.049 (.076)	.133** (.052)	.301* (.156)	.115 (.078)
Treatment int	.010 (.047)	.078* (.044)	-.068 (.054)	-.042 (.048)	-.023 (.061)	.311 (.250)	.067 (.108)	-.100 (.068)	-.606** (.261)	-.084 (.102)
Obs.	2371	2371	2371	2371	2371	2371	2371	2371	2371	2371
Panel B: Baseline age										
Treatment	-.081 (.104)	-.029 (.094)	.211* (.123)	-.027 (.103)	.100 (.134)	-.612 (.455)	-.263 (.238)	.035 (.130)	-.963 (.657)	.042 (.256)
Treatment int	.014 (.009)	.004 (.009)	-.011 (.011)	.006 (.009)	-.003 (.012)	.027 (.040)	.023 (.022)	.005 (.012)	.086 (.059)	.003 (.023)
Obs.	2371	2371	2371	2371	2371	2371	2371	2371	2371	2371
Panel C: Maternal education										
Treatment	.071*** (.025)	.007 (.026)	.108*** (.029)	.039 (.025)	.061* (.033)	-.331** (.132)	-.011 (.058)	.095*** (.035)	-.119 (.147)	.080 (.052)
Treatment int	-.003 (.054)	.077* (.046)	-.074 (.045)	.017 (.042)	.067 (.063)	.170 (.168)	.025 (.088)	-.017 (.054)	.510* (.276)	-.020 (.100)
Obs.	2371	2371	2371	2371	2371	2371	2371	2371	2371	2371

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent (i.e. higher values of empowerment index components all correspond to higher levels of empowerment). Marital expectations index is not mean 0 because married girls are assigned the minimum value calculated for non-married girls. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A17: Non-cognitive Skills, Survey Measures: Heterogeneous Effects for Household Shocks

	Socio-emotional index (1)	Freedom of movement index (2)	Empowerment index (3)	Self-esteem index (4)	Future planning index (5)	Marital expectations index (6)	Educ. / emp. aspirations index (7)	Gender norms index (8)	Cantril's ladder (9)	Enumerator assessment index (10)
Panel A: Economic shock										
Treatment	.080*** (.031)	.032 (.035)	.085** (.034)	.005 (.034)	.067 (.044)	-.328** (.153)	.026 (.066)	.127*** (.038)	-.007 (.176)	.067 (.067)
Treatment int	-.018 (.033)	-.022 (.040)	.014 (.036)	.060 (.042)	.006 (.052)	.017 (.136)	-.064 (.066)	-.065 (.045)	-.040 (.179)	.007 (.080)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380
Panel B: Crime shock										
Treatment	.079*** (.026)	.019 (.024)	.089*** (.028)	.042 (.026)	.066** (.033)	-.337*** (.126)	.009 (.058)	.090*** (.034)	-.023 (.129)	.092* (.054)
Treatment int	-.071 (.058)	-.0003 (.048)	.031 (.057)	-.009 (.051)	.026 (.065)	.182 (.217)	-.157* (.091)	-.0006 (.066)	-.044 (.298)	-.143 (.092)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380
Panel C: Death/illness shock										
Treatment	.069** (.027)	.022 (.028)	.116*** (.030)	.048* (.027)	.095*** (.037)	-.298** (.134)	-.005 (.063)	.083** (.039)	.015 (.149)	.076 (.058)
Treatment int	.002 (.039)	-.005 (.040)	-.055 (.034)	-.019 (.035)	-.062 (.048)	-.045 (.122)	-.017 (.061)	.015 (.044)	-.101 (.184)	-.006 (.065)
Obs.	2380	2380	2380	2380	2380	2380	2380	2380	2380	2380

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent (i.e. higher values of empowerment index components all correspond to higher levels of empowerment). Marital expectations index is not mean 0 because married girls are assigned the minimum value calculated for non-married girls. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A18: Non-cognitive Skills - Demonstration Tasks and Endline Psycho-Social Indices: Heterogeneous Effects

	Delay discounting (1)	Completed mirror drawings (2)	Mirror drawings (seconds) (3)	Scavenger hunt index (4)	Locus of control control index (5)	Perceived stress index (6)	Rosenberg self-esteem index (7)
Panel A: School quality							
Treatment	.021 (.040)	-.063 (.110)	-4.474 (6.389)	-.100 (.078)	-.019 (.071)	-.035 (.065)	.063 (.045)
Treatment int	-.050 (.064)	.214 (.171)	13.303 (8.974)	.035 (.115)	.014 (.094)	.031 (.094)	-.090 (.059)
Obs.	2344	2351	2281	2344	2344	2344	2344
Panel B: Baseline age							
Treatment	.120 (.122)	.412 (.362)	7.366 (21.196)	-.002 (.208)	-.119 (.212)	.064 (.201)	-.048 (.121)
Treatment int	-.011 (.011)	-.033 (.031)	-.491 (1.892)	-.007 (.018)	.009 (.019)	-.008 (.018)	.006 (.011)
Obs.	2378	2385	2315	2378	2378	2378	2378
Panel C: Maternal education							
Treatment	.004 (.032)	.030 (.079)	2.783 (4.752)	-.080 (.060)	-.002 (.049)	.003 (.050)	.011 (.030)
Treatment int	-.032 (.046)	.164 (.166)	-1.638 (9.727)	.073 (.123)	-.123 (.104)	-.197* (.115)	.028 (.058)
Obs.	2371	2378	2308	2371	2371	2371	2371

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. For Columns 5-7 measures added at endline, we control for lagged values of overall life skills indices. This specification was pre-specified.

Delay discounting is an indicator for whether the respondent would prefer 60 Rs. in one week over 30 Rs. now (respondents were informed that they would have a chance to receive a gift valued correspondingly). Completed mirror drawings takes on values from 0 to 4 and Mirror drawings (seconds) measures the total number of seconds spent on mirror drawings, conditional on having attempted at least one mirror drawing.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A19: Non-cognitive Skills - Demonstration Tasks and Endline Psycho-Social Indices: Heterogeneous Effects for Household Shocks

	Delay discounting (1)	Completed mirror drawings (2)	Mirror drawings (seconds) (3)	Scavenger hunt index (4)	Locus of control control index (5)	Perceived stress index (6)	Rosenberg self-esteem index (7)
Panel A: Economic shock							
Treatment	.042 (.040)	.068 (.101)	.243 (5.993)	-.025 (.068)	-.016 (.064)	-.028 (.079)	.004 (.040)
Treatment int	-.070* (.039)	-.021 (.093)	3.251 (6.749)	-.088 (.077)	-.004 (.089)	.005 (.097)	.020 (.042)
Obs.	2380	2387	2317	2380	2380	2380	2380
Panel B: Crime shock							
Treatment	.0001 (.034)	.052 (.086)	3.489 (4.779)	-.082 (.059)	-.039 (.048)	-.035 (.052)	.018 (.030)
Treatment int	-.006 (.053)	.033 (.143)	-9.940 (9.638)	.033 (.120)	.174 (.129)	.079 (.122)	-.012 (.053)
Obs.	2380	2387	2317	2380	2380	2380	2380
Panel C: Death/illness shock							
Treatment	.022 (.033)	.073 (.091)	5.299 (5.753)	-.046 (.064)	-.035 (.056)	-.014 (.056)	.013 (.034)
Treatment int	-.055* (.033)	-.040 (.105)	-7.797 (7.953)	-.082 (.081)	.048 (.079)	-.025 (.086)	.006 (.038)
Obs.	2380	2387	2317	2380	2380	2380	2380

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. For Columns 5-7 measures added at endline, we control for lagged values of overall life skills indices. This specification was pre-specified.

Delay discounting is an indicator for whether the respondent would prefer 60 Rs. in one week over 30 Rs. now (respondents were informed that they would have a chance to receive a gift valued correspondingly). Completed mirror drawings takes on values from 0 to 4 and Mirror drawings (seconds) measures the total number of seconds spent on mirror drawings, conditional on having attempted at least one mirror drawing.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A20: Non-cognitive Skills, Parental Reports: Heterogeneous Effects

	Parental perception of girl's strengths (1)	Parental perception of girl's self-efficacy (2)	Parental perception of freedom of movement (3)	Parent daughter communication (4)	Parental gender attitudes (5)	Parental schooling attitudes (6)	Parental marriage attitudes (7)
Panel A: School quality							
Treatment	-0.13 (.025)	.029 (.043)	.037 (.037)	.054 (.043)	.039 (.037)	.075 (.054)	.067 (.045)
Treatment int	-.056 (.037)	-.054 (.059)	-.045 (.057)	-.135** (.057)	-.079 (.053)	-.094 (.084)	-.085 (.063)
Obs.	2398	2394	2398	2398	2398	2398	2398
Panel B: Baseline age							
Treatment	-.096 (.091)	-.088 (.137)	.197 (.142)	-.050 (.130)	.050 (.101)	-.128 (.224)	-.370** (.183)
Treatment int	.005 (.008)	.008 (.013)	-.016 (.013)	.003 (.012)	-.005 (.009)	.014 (.021)	.036** (.017)
Obs.	2432	2428	2432	2432	2432	2432	2432
Panel C: Maternal education							
Treatment	-.048** (.020)	.007 (.032)	.015 (.031)	-.024 (.029)	.013 (.027)	.043 (.043)	.026 (.034)
Treatment int	.048 (.044)	.024 (.061)	.045 (.049)	.077 (.048)	-.046 (.052)	-.010 (.074)	.004 (.067)
Obs.	2426	2422	2426	2426	2426	2426	2426

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A21: Non-cognitive Skills, Parental Reports: Heterogeneous Effects for Household Shocks

	Parental perception of girl's strengths (1)	Parental perception of girl's self-efficacy (2)	Parental perception of freedom of movement (3)	Parent daughter communication (4)	Parental gender attitudes (5)	Parental schooling attitudes (6)	Parental marriage attitudes (7)
Panel A: Economic shock							
Treatment	-0.056** (.028)	-0.022 (.040)	.033 (.039)	-0.015 (.037)	-0.005 (.032)	.078 (.054)	.061 (.044)
Treatment int	.025 (.029)	.044 (.051)	-.020 (.044)	.001 (.037)	.008 (.034)	-.077 (.054)	-.064 (.050)
Obs.	2434	2430	2434	2434	2434	2434	2434
Panel B: Crime shock							
Treatment	-0.033* (.020)	.004 (.030)	.014 (.031)	-.011 (.029)	.006 (.028)	.019 (.046)	.021 (.034)
Treatment int	-.069 (.044)	.002 (.076)	.056 (.061)	-.027 (.049)	-.043 (.048)	.097 (.086)	.013 (.063)
Obs.	2434	2430	2434	2434	2434	2434	2434
Panel C: Death/illness shock							
Treatment	-.023 (.022)	.023 (.037)	.008 (.033)	-.007 (.031)	-.007 (.031)	.038 (.052)	.029 (.039)
Treatment int	-.048* (.028)	-.046 (.047)	.034 (.042)	-.019 (.036)	.016 (.035)	-.015 (.059)	-.018 (.043)
Obs.	2434	2430	2434	2434	2434	2434	2434

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

For all included indices, we first take the difference between each component survey response value and the mean within the control group and then divide by the control group standard deviation. We then average over all index components, ensuring that values for each component are constructed so that the index interpretation is consistent. Detailed definitions of all referenced indices can be found in the analysis plan posted on-line.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A22: Child Labor: Heterogeneous Effects

Married	Child works (Economically active)	Child works for pay	Child works outside of family	Child labor	Hazardous child labor	Other worst forms of child labor	Hours worked in a day	Hours worked unpaid work	Hours active (Paid + unpaid)	Hour active outside house
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: School quality										
Treatment	.090** (.045)	.041 (.031)	.026 (.041)	.015 (.053)	.008 (.047)	-.010 (.031)	.159 (.184)	-.065 (.112)	.094 (.232)	.055 (.108)
Treatment int	-.092 (.059)	-.038 (.051)	-.067 (.060)	-.014 (.075)	.007 (.073)	.059 (.043)	-.163 (.277)	.205 (.148)	.042 (.344)	-.096 (.173)
Obs.	2399	2350	2351	2350	2350	2351	2350	2350	2350	2350
Panel B: Baseline age										
Treatment	.068 (.117)	.101 (.119)	-.022 (.101)	-.049 (.128)	-.036 (.123)	.018 (.085)	.689 (.594)	.696 (.516)	1.385 (.910)	.683* (.351)
Treatment int	-.002 (.010)	-.007 (.011)	.001 (.009)	.005 (.011)	.004 (.011)	.0003 (.008)	-.056 (.060)	-.061 (.048)	-.118 (.086)	-.062* (.036)
Obs.	2433	2384	2385	2384	2384	2385	2384	2384	2384	2384
Panel C: Maternal education										
Treatment	.041 (.031)	.039 (.039)	-.026 (.030)	-.009 (.036)	-.007 (.037)	.026 (.023)	.066 (.158)	.029 (.078)	.095 (.192)	-.007 (.097)
Treatment int	-.021 (.036)	.040 (.059)	.090* (.052)	.066 (.056)	.074 (.054)	-.033 (.040)	-.109 (.208)	-.065 (.146)	-.174 (.279)	-.003 (.132)
Obs.	2426	2377	2378	2377	2377	2378	2377	2377	2377	2377

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

Married is an indicator variable for whether girl is married or committed (engaged). The set of survey questions used to construct each of the indicator variable outcomes in Columns (2) through (7) can be found in the analysis plan posted on-line. Time use outcomes in Columns (8) through (11) are defined based on time use patterns recorded for “a typical day in the past week.”

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A23: Child Labor: Heterogeneous Effects for Household Shocks

	Married	Child works (Economically active)	Child works for pay	Child works outside of family	Child labor	Hazardous child labor	Other worst forms of child labor	Hours worked in a day	Hours worked unpaid work	Hours active (Paid + unpaid)	Hour active outside house
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treatment	.041 (.036)	.023 (.047)	.020 (.031)	-.037 (.035)	-.031 (.046)	-.026 (.046)	.011 (.032)	.045 (.173)	-.019 (.105)	.026 (.217)	-.002 (.100)
Treatment int	.002 (.034)	.043 (.043)	.001 (.034)	.043 (.033)	.058 (.044)	.058 (.045)	.016 (.037)	.025 (.160)	.074 (.117)	.100 (.202)	.0003 (.121)
Obs.	2435	2386	2386	2387	2386	2386	2387	2386	2386	2386	2386
Panel A: Economic shock											
Treatment	.048 (.030)	.050 (.040)	.018 (.027)	-.017 (.030)	.001 (.038)	.009 (.038)	.015 (.022)	.084 (.147)	.012 (.078)	.096 (.186)	.008 (.094)
Treatment int	-.051 (.050)	-.007 (.057)	.030 (.050)	.045 (.043)	.018 (.058)	-.001 (.062)	.043 (.054)	-.181 (.270)	.096 (.167)	-.085 (.337)	-.049 (.157)
Obs.	2435	2386	2386	2387	2386	2386	2387	2386	2386	2386	2386
Panel B: Crime shock											
Treatment	.015 (.031)	.029 (.041)	.005 (.030)	-.041 (.032)	-.016 (.038)	-.015 (.039)	.028 (.023)	-.012 (.165)	-.014 (.091)	-.026 (.208)	-.012 (.111)
Treatment int	.068** (.031)	.049 (.041)	.040 (.031)	.075** (.035)	.049 (.043)	.059 (.042)	-.019 (.030)	.177 (.176)	.096 (.123)	.274 (.234)	.030 (.127)
Obs.	2435	2386	2386	2387	2386	2386	2387	2386	2386	2386	2386
Panel C: Death/illness shock											

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. This specification was pre-specified.

Married is an indicator variable for whether girl is married or committed (engaged). The set of survey questions used to construct each of the indicator variable outcomes in Columns (2) through (7) can be found in the analysis plan posted on-line. Time use outcomes in Columns (8) through (11) are defined based on time use patterns recorded for “a typical day in the past week.”

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A24: Cognitive Skills: Heterogeneous Effects

	Survey data				Administrative data			
	Hours studying at home (1)	Total hours spent at school (2)	ASER score Mathematics (3)	ASER score Hindi (4)	ASER score English (5)	GPA Grade 6 (6)	GPA Grade 7 (7)	GPA Grade 8 (8)
Treatment	-.133 (.102)	-.116 (.234)	-.016 (.120)	-.097 (.149)	-.022 (.135)	-.066 (.113)	.088 (.104)	.136 (.105)
Treatment int	.135 (.154)	.536 (.376)	-.008 (.156)	.224 (.189)	-.114 (.183)	-.103 (.150)	-.404** (.179)	-.336** (.162)
Obs.	2350	2350	2344	2344	2344	2144	1946	1884
			Panel A: School quality					
Treatment	.006 (.418)	-.847 (1.089)	-.124 (.263)	.246 (.324)	-.045 (.320)	.001 (.253)	-.135 (.271)	.184 (.221)
Treatment int	-.007 (.037)	.093 (.100)	.010 (.023)	-.020 (.029)	-.003 (.029)	-.011 (.022)	.001 (.021)	-.020 (.018)
Obs.	2384	2384	2378	2378	2378	2176	1974	1910
			Panel B: Baseline age					
Treatment	-.043 (.085)	.149 (.212)	-.034 (.079)	.039 (.103)	-.089 (.093)	-.141* (.076)	-.120 (.091)	-.061 (.085)
Treatment int	-.069 (.161)	.250 (.337)	.125 (.122)	.039 (.137)	.161 (.125)	.127 (.106)	.004 (.122)	.177* (.105)
Obs.	2377	2377	2371	2371	2371	2152	1956	1896
			Panel C: Maternal education					

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. In Columns 3-5, controls for baseline outcome values cannot be included since cognitive tests were not conducted at baseline; we instead include controls for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school, and grades as reported in grade five. This specification was pre-specified.

Time use outcomes in Columns (1) and (2) are defined based on time use patterns recorded for “a typical day in the past week.” ASER test score outcomes in Columns (3) through (5) and GPA outcomes in Columns (6) through (8) take on values between zero and four.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table A25: Cognitive Skills: Heterogeneous Effects for Household Shocks

	Survey data				Administrative data			
	Hours studying at home (1)	Total hours spent at school (2)	ASER score Mathematics (3)	ASER score Hindi (4)	ASER score English (5)	GPA Grade 6 (6)	GPA Grade 7 (7)	GPA Grade 8 (8)
Treatment	-0.15 (.101)	.282 (.222)	-.104 (.089)	.0002 (.108)	-.095 (.105)	-.106 (.080)	-.181 (.116)	.041 (.096)
Treatment int	-.079 (.115)	-.164 (.254)	.140 (.090)	.054 (.118)	.038 (.109)	-.021 (.069)	.102 (.087)	-.123 (.075)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912
	Panel A: Economic shock							
Treatment	-.069 (.078)	.162 (.194)	-.007 (.080)	.059 (.097)	-.028 (.094)	-.099 (.074)	-.093 (.094)	-.0006 (.083)
Treatment int	.056 (.145)	.153 (.328)	-.111 (.115)	-.207 (.152)	-.349** (.151)	-.147 (.090)	-.196* (.113)	-.255** (.100)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912
	Panel B: Crime shock							
	Panel C: Death/illness shock							
Treatment	-.003 (.088)	.302 (.240)	-.028 (.084)	.108 (.104)	-.021 (.093)	-.096 (.074)	-.089 (.094)	-.031 (.088)
Treatment int	-.147 (.112)	-.294 (.254)	.018 (.079)	-.189* (.109)	-.132 (.093)	-.056 (.074)	-.077 (.068)	-.006 (.063)
Obs.	2386	2386	2380	2380	2380	2178	1976	1912

Notes: Table contains results from regressing the outcome variable indicated by the column header on an indicator for treatment (reported), the interaction of treatment with the specified characteristics (reported), the characteristics, stratification fixed effects, age fixed effects, baseline value of the outcome, and a vector of dummies for the most important type of employment in the household at baseline. In Columns 3-5, controls for baseline outcome values cannot be included since cognitive tests were not conducted at baseline; we instead include controls for baseline school dropout status, attendance, grade progression, time spent studying, hours spent on school, and grades as reported in grade five. This specification was pre-specified.

Time use outcomes in Columns (1) and (2) are defined based on time use patterns recorded for “a typical day in the past week.” ASER test score outcomes in Columns (3) through (5) and GPA outcomes in Columns (6) through (8) take on values between zero and four.

Standard errors, clustered by school, in parenthesis.

* significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.