Effects of speed-schools in Niger*

Anne Kielland[1], Andreas Kotsadam[2], and Jing Liu[1]

[1]FAFO
[2]Ragnar Frisch Centre for Economic Research

June 6, 2024

Abstract

We evaluate a two-year accelerated education program in Niger by randomizing schools to villages and places in the schools within villages. The program had large effects on education and learning: 183 percent more treated children start lower secondary education (our main educational outcome), they are 121 percent more likely to be literate, and they score 75 percent higher on a numeracy index. We find no indications of spillover effects of the program. In particular, control children in treated villages are very similar to children in control villages at endline. Despite the large effects on educational and learning outcomes we find no effects on any of our other five main pre-registered outcomes: Attitudes toward the appropriate marriage age and gender equality, well-being, religious extremism, or engaging in hazardous work.

*Funding from the Norwegian Research Council is acknowledged. An analysis plan is pre-registered at the AEA RCT registry (AEARCTR-0006878) and we present all main results exactly as specified there unless explicitly noted in the text. Exploratory analyses outlined in the plan that are not presented here are presented in a separate document (a populated pre-plan). The pre-analysis plan is added as an Appendix and the populated pre-plan can be found at https://andreaskotsadam.wordpress.com/wp-content/uploads/2024/06/populated-pap-ssa-niger.pdf. Anonymized replication files will be available upon acceptance.

1 Introduction

While school attendance has dramatically improved in many of the world's poorest countries over recent decades, unfortunately, this has not translated into large learning gains. Many children, despite years of schooling, remain illiterate and innumerate (Glewwe and Muralidharan, 2016). The causes of this "learning crisis" are multifaceted, ranging from a low demand for education and insufficient educational resources, to a shortage of competent teachers (World Bank, 2018; Kremer et al., 2013; Pritchett, 2013; Fazzio et al., 2021). The situation is particularly grim in Niger, where less than half of primary school-aged children are enrolled in schools. Moreover, the quality of learning is poor, with less than 10 percent of sixth graders possessing reading and counting skills (World Bank, 2018). This predicament has left millions of older children behind, a situation that needs urgent rectification even if attendance and learning rates for younger children were to improve.

One approach to reintegrating older children, who have either never enrolled, dropped out, or did not learn enough in primary school to qualify for secondary school, involves admitting them into 'speed-schools' or accelerated education programs. These programs vary in form but universally feature condensed schooling, with the aim of preparing learners for advanced educational stages in less time. We are studying a two-year speed-school program in Niger that caters to rural out-of-school children aged between 12 and 14, who have never enrolled or prematurely left primary education without qualifying for further education. The program covers the first six years of primary school curriculum within two years, intending to provide these children with a second chance to catch up and reintegrate into formal schools.

We are closely collaborating with the Strømme Foundation, an NGO spearheading speed-schools in the Sahel region, to study the impacts of their speed-schools using a two-stage randomization procedure. We administer treatment randomization at both the community (across 85 communities) and individual levels (across 1,786 children in treated communities). This process creates 'treatment' and 'control' communities, as well as 'treatment', 'spillover', and 'pure control' children within these communities.

We find substantial effects on both education and learning outcomes around 2.5 years after the intervention started. Whereas only 12 percent of the control group children in the treatment communities advance to lower secondary education, this number increases to 34 percent for the treated children. The offer to attend speed-school led to a remarkable 121 percent increase in literacy rates, rising from 14 to 31 percentage points. Furthermore, numeracy skills were enhanced by 0.44 standard deviations, which corresponds to a 75 percent increase from the control group mean. The impact on learning was so pronounced that children who were offered two years of speed-school education on average demonstrated superior skills in mathematics, reading, and writing than those in the control group with six years of primary education. After adjusting for non-perfect compliance and calculating the average treatment effects on the treated, the results appear even more striking. The learning effects of the intervention are among the top 10 percent of outcomes from education interventions in low and middle-income countries (Evans and Yuan, 2022). We did not observe any negative or positive spillover effects on either schooling or learning within the treated communities. We find clear treatment effects for all levels of previous education of the children and no statistically significant heterogeneity.

Our primary contribution is to the literature on enhancing educational outcomes for children in impoverished settings. Specifically, speed-schools provide an avenue to catch up and reintegrate into the formal education system for children who would otherwise have few other opportunities. While speed-schools are becoming increasingly prevalent in Sub-Saharan Africa, there are very few rigorous evaluations of these initiatives. Two notable exceptions are Akyeampong et al. (2018) and Diazgranados-Ferráns et al. (2022). Akyeampong et al. (2018) evaluate a 10 month speed school program in Ethiopia that targets children aged 9-14. They have longitudinal data over 6 years and compare children in the speed school with children in other schools. They find that attendance and learning were higher for the speed school children. There are, however, also other concerning differences over time for the different groups of children. For instance, the relative household assets and livestock improvements for the speed school children households were very large during this period. Diazgranados-Ferráns et al. (2022) investigate the effects of a nine month speed-school on 1700 children aged between 9 and 14 in Nigeria. The study design was an individual-level RCT, but the low adherence to the assigned treatment conditions resulted in an imbalance in the samples based on e.g., prior education levels. Controlling for baseline imbalances, in particular in a difference-in-differences analysis, revealed positive program effects on both numeracy and literacy, albeit smaller than the effects we observe. We contribute to this small literature by investigating a longer program as well as by investigating spillover effects and effects on other important outcomes. Measuring spillover effects is vital for ensuring the accuracy and validity of the findings, understanding the full impact of the intervention, and for ethical reasons as some interventions might inadvertently harm non-target individuals.

We also contribute to the literature on remedial education in developing countries. Banerjee et al. (2007) evaluated the Pratham model, which, similar to the speed-school program we study, employs educated individuals from the local community and ensures a low student-teacher ratio to provide teaching at an appropriate level. A notable distinction, however, is that this model targets children who are lagging behind in their current studies, as opposed to our focus on out-of-school children. Another related remedial education program is the one in Niger studied by Aker and Ksoll (2019), who focused on adult education with participants having a mean age of 35 years old. In a cluster RCT at the village level, they found that the program increased math and reading test scores by around 0.2 standard deviations.

Previous literature demonstrates that targeted instruction and structured lesson plans, both integral components of our program, are extremely effective in low-income contexts

(Angrist and Meager, 2023; Angrist et al., 2020), albeit expensive. Our intervention costs the NGO roughly 350 USD per child per year, a figure that can be compared to the average annual per pupil expenditure of 279 USD in Sub-Saharan Africa at large. However, the learning gains achieved are substantial. Two other recent successful interventions in Guinea Bissau (Fazzio et al., 2021) and Gambia (Eble et al., 2021), which report even larger learning improvements, had costs ranging from 240 USD in Gambia (supplementary to regular school costs as it was an after-school program) to over 400 USD in Guinea Bissau. Quality education indeed comes with a price tag, but these and our findings highlight that it is possible to accomplish significant advancements. Future studies should focus on trying to reduce the costs of successful programs.

In addition to the results on educational outcomes we also test whether the program had any effects on other variables that are often hypothesized to be affected by education (such as outcomes related to gender equality, early marriage, and well-being) or that are particularly salient in our context (such as religious extremism or hazardous child work). We find very limited effects on these outcomes.

2 Context and intervention

Located in West Africa, Niger is a Francophone country with school enrollment and completion rates among the lowest worldwide. Although households in Niger can formally enroll their children in primary school without charge, the reality often includes some ancillary school-related expenditures, in addition to the opportunity costs of children's time (Bagby et al., 2016). While education is officially mandatory for children aged 7 to 12, this law has not been consistently enforced, especially in rural areas. Niger faces challenges with education due to a rapidly increasing population growth (3.7%) combined with erratic and slow economic growth, making it difficult to train more teachers and build more schools (World Bank, 2023).

Niger has long held the unfortunate last position among the 189 countries in the UNDP Human Development Index, with education being a significant contributing factor to its low ranking. Dosso and Tillaberi, the regions targeted by our intervention study, suffer from extreme deprivation, characterized by a child mortality rate of 123 per 1000, stunting rates of 47 percent, and acute malnutrition affecting almost 11 percent of children under five (DHS, 2012; INS, 2022, 2023).

The public education system, which promised parents an escape route from poverty, has widely failed to deliver quality education, resulting in widespread disillusionment—a phenomenon referred to as "l'aversion scolaire" (Moussa, Forthcoming). Formal schools' learning outcomes are disheartening, with very few students exiting primary school having achieved minimum proficiency in reading and mathematics (UNESCO, 2021). In such a context, encouraging early teenagers to return to school is challenging.

The intervention we are studying offers "speed-schools" or accelerated learning programs, specifically catering to out-of-school children who surpass the usual primary school age. These speed-schools aim to bridge the education gap for these children by providing a condensed form of the primary curriculum. The particular program we study, termed Stratégie de Scolarisation Accélérée 2 (SSA2), distills the content of grades 1-6 into an intensive two-year course. The focus of the pedagogy is to have small enough classes to enable individually adapted teaching at the correct level. Classes are held 6 days a week. The total teaching load amounts to 34 hours per week, out of which 24 hours focus explicitly on Math and French and the rest on Social Science, Science, and Arts. As opposed to in the regular schools, where teaching is in French only, teaching in other subjects in the speed schools starts in the mother tongue for the first two months before moving over to French only.

The SSA2 model of accelerated learning is designed to be sustainable and scalable within the context of a low-income economy. The program has been designed to avoid creating a parallel structure to the existing formal school system, which could potentially result in children opting out of public schooling. To prevent children from dropping out of formal school to join the speed-school, the program is not run consecutively in the same community.

The program employs an "animator" instead of a licensed teacher, deliberately avoiding draining valuable teacher resources from public schools. These animators are ideally local individuals who have reached the tenth grade of the formal education system. This approach taps into under-utilized local intellectual resources. The animators first receive one month of initial training before the first year and a one week refresher training the following year. There are also three shorter pedagogical meetings per year focusing on capacity building. The animators are paid around 50 percent more than contractual teachers in regular schools, but the work is temporary in nature as the school moves after two years. There are few problems of teacher absenteeism as teachers are monitored by community members and by the Strømme foundation conducting random visits.

The program tries to minimize alternative costs to society. The classrooms are makeshift structures constructed from local materials like branches, built by the local population. These classrooms can be dismantled during the rainy season and reconstructed for the second year of the program, with the materials being stored and recycled, thereby avoiding the need for permanent, solid structures that could be otherwise utilized. This helps in keeping capital investments minimal.

The main costs of SSA2 are in monitoring and supervision structures, as travel is time and gas-intensive in rural areas. The Global Education Monitoring Report estimates that the per-pupil and year cost for primary school children in Sub-Saharan Africa is estimated to be 279 USD. Given that SSA2 targets older and particularly disadvantaged children in rural areas the costs were expected to be higher and are estimated to be 350 USD per year. It is, however, difficult to know how to compare an intervention that is supposed to cover six years with numbers for single years of non-quality education.

The SSA2 project model has several objectives, each designed to address specific social and cultural challenges in the region:

1. **Boosting Education and Learning:** The primary goal of SSA2 is to enhance educational outcomes and increase enrollment in lower secondary education, particularly emphasizing on girls' education.

2. **Reducing Early Marriages and Pregnancy:** The program aims to address the prevalent issue of early marriages and pregnancies among girls. According to DHS (2012), nearly 80 percent of girls in Niger are married before the age of 18, and 24 percent before the age of 15. Secondary school attendance has been shown to reduce female fertility in Ghana and Kenya (Duflo et al., 2021, 2015) and it is possible that this primary school intervention affects perceptions of the appropriate marriage age.

3. **Reducing Hazardous Child Labor:** SSA2 seeks to curb child labor, which, as per our data, implicates approximately 20 percent of the children in the control group.

4. **Preventing Youth Recruitment into Violent Groups:** SSA2 intends to provide an alternative path for youth and reduce their vulnerability to recruitment by violent groups prevalent in the region. Niger, along with neighboring countries, has witnessed a surge in armed attacks often initiated by misguided individuals invoking religious motives, frequently labelled as jihadists.

   In addition to these explicit goals, we also add the following goals:

5. **Improving Gender Attitudes:** The program aims to foster more balanced gender attitudes, as over 60 percent of children in control communities concur with the idea that it is more important to send boys to school than girls.

6. **Enhancing General Wellbeing:** By creating access to quality education, the

program seeks to improve the overall wellbeing and self-esteem of children.

3 Randomization and sampling

It is important to rigorously evaluate educational interventions, particularly in the context of the ongoing learning crisis. We posit that randomization is ethically justified, as it does not reduce the available spots in schools, but rather ensures children of comparable need are granted an equal opportunity for program participation. Ex-ante, it was furthermore not certain that SSA2 would be successful in teaching valuable skills to the children and it was unsure whether the benefits exceeded the opportunity costs.

We conducted a listing exercise to map the need in different communities in the two regions. Children in 120 communities were listed based on the following criteria: household within what is locally described as the "academic distance" from the village center/prospective Speed-school location (5 km), age (around 12-14), out of school, willingness to go to school, parental support and consent. Note that the listing did not include the number of years of previous school attendance. As a result we have around 10 percent of the sample that state that they have more than six years of education (see e.g. Appendix Figure A.2). The most likely reason for these children being out of school is that they did not qualify for secondary school and that the years also include grade repetition and non-formal schools. Unfortunately, we do not have any baseline test of competency to see their skills at baseline but we note that the literacy and numeracy at endline is low for the control group.

Based on the lists we selected 85 communities that had sufficiently many children and carried out the randomization. The experiment is randomized at two different levels:

Randomization across communities: Our experiment takes place in 85 communities in two regions. 30 communities were randomized to treatment and 55 to control. It was decided that there would be 15 treated communities in each of two regions covered (Dosso and Tillaberi) and therefore randomization to treatment and control communities were done within regions. We show the location of the communities in Appendix Figure A.1. In the control communities we surveyed 1,430 children at baseline.

Randomization across individuals: In the treatment areas we had larger samples to allow for comparisons of individual treatment and control groups. We surveyed 1,786 children at baseline. The schools want around 28 individuals per school and since treatment compliance was expected to be imperfect we randomized around 35 individuals to treatment. Randomization to the program at the individual level was stratified by gender.

We collected baseline data in August to November 2018. Baseline interviews in the treatment communities were blinded and treatment status was revealed after data collection in the village was finished. Treatment started shortly after baseline data were collected. The intervention was completed in June 2020 and the second round of data collection took place during December 2020-June 2021.

4 Main outcomes

A goal of the evaluation is to provide a comprehensive picture of the impacts of speed-schools. We therefore collect data on a large number of outcomes and we report the full breadth of the evidence. That being said, power considerations forced us to limit the number of main hypotheses tested. We group the outcomes into 6 groups and for each group we have either an index or a main variable to be tested. For each topic we also registered a number of secondary analyses that can be seen in the pre-analysis plan (added to the Appendix). We only present the results of these secondary analyses in the main paper if there are effects on the main outcomes.[1] We chose the following main groups with the corresponding main variables in italics:

[1]We follow the suggestions of Banerjee et al. (2020) and present the full set of results in a populated pre-analysis plan, available at https://andreaskotsadam.wordpress.com/wp-content/uploads/2024/06/populated-pap-ssa-niger.pdf

1. Schooling and learning outcomes. *Starting lower-secondary education.*
2. Marriage and fertility. *Appropriate marriage age.*
3. Support for violence. An index of *Religious extremism.*
4. Self-esteem and wellbeing. A *Rosenberg index* of self-esteem.
5. Hazardous child labor and time use. *Hazardous child labor.*
6. Gender attitudes. A *Gender equality index.*

We describe the coding of the variables in detail in Appendix Section A.1 but we here give an overview of the main variables. The overarching program goal of SSA2 is for out-of-school children to transfer back into lower secondary level of the traditional schooling system in Niger. Our main measure for the educational outcomes is *Starting lower-secondary education.* We also analyze a number of secondary variables for education and learning such as *Qualified for secondary*, *Years of schooling*, and *Highest grade level.*

We have several variables intended to measure learning. *Illiterate* is a measure based on self reported illiteracy and, if the child answers that they can read and write, we still code them as illiterate if they cannot write their own name on a piece of paper we provide them or if they cannot read the sentence "Fatima likes going to school, but she also likes to help her father". The sentence was invented by the research team and did not feature in any education material at the schools. *Numeracy index* ranges from 0 to 4 where one point is given for being able to solve each of the problems we give them on addition (77+11= ?), subtraction (49-12= ?), division (28/4= ?), and multiplication (4x5= ?). As with the literacy question, the teachers did not know what we asked the children. The other main variables are described in enumerated list above and details are provided in the Appendix.

5 Empirical strategy

Our main empirical strategy is intended to capture reduced form (intention to treat) individual treatment effects, spillover effects, and effects at the community level. In

Section 6.2 we also present treatment effects on the treated by means of instrumental variables estimation.

5.1 Individual level effects

For the analysis of individual level treatment effects we restrict the sample to the treated communities and estimate the following regression:

$$Y_{2i} = \beta Treated_i + \lambda X_{1i} + \varepsilon_i \tag{1}$$

where 2 indicates follow up and 1 indicates baseline, i indexes individuals. The vector X always includes fixed effects for the strata variables gender and community. In addition we investigate if we can improve precision in the estimates by including the controls listed in Table 1 and by picking optimal controls from the total list of controls using LASSO (Belloni et al., 2014).[2] If we have missing values on explanatory variables we code the variables as zero and include dummy variables controlling for missing status so that we do not lose observations. We use robust standard errors in all estimations unless otherwise stated. As there is no general imbalance across treatment and control (see below), the estimation with only the strata variables is our main specification as pre-specified.

5.2 Spillovers and community level effects

The specification in (1) is a biased estimate of the overall effect in a community if there are spillover effects within villages. The question of spillovers is also interesting in itself. To estimate spillover effects we restrict the sample to control individuals and estimate the following regression:

[2]See Appendix Section A.1 for coding of the variables. As there is no limitation on how many variables that can meaningfully be included in the LASSO regressions we also include baseline variables and household variables for the variables listed in Appendix Section A.1 when they exist. In addition we explore whether the variables on building material for the roof, floor, and walls help us in precision beyond the asset index. In total we include 43 covariates and their interactions.

$$Y_{2ic} = \beta TreatedCommunity_{ic} + \lambda X_{1i} + \varepsilon_{ic} \quad (2)$$

where the treatment community is randomly assigned at the community level (c). As treatment at the community level is randomly assigned within regions we always include Region as a strata variable in these regressions. We still estimate this regression using individual level data as power may be increased by including individual level controls. We do, however, cluster the standard errors at the community level (c) in these regressions. As the sample includes only non-treatment individuals, β identifies within-community spillover effects by comparing control individuals in treatment villages to control individuals in pure control villages. Note that we have not randomly assigned the treatment share in different communities so we can not investigate how the spillovers change with different intensities of the treatment without further assumptions (Baird et al., 2018). We also use specification (2) but without excluding any individuals to estimate the total effect of having a speed-school in the area.[3]

5.3 Descriptive statistics and balance tests

In Table 1 we show descriptive statistics for the whole sample at endline and the different treatment and control groups. We see that one third of all the children are treated and that the share of treated in the treatment communities is 60 percent. While there are 30 treated communities and 55 control communities, 57 percent of the surveyed children live in treatment communities as we interview more children there to allow for individual level comparisons. The share of girls in the sample is 47 percent.

We have six main outcome variables. As these are measured at endline we expect there to be differences across samples if there are treatment (and spillover) effects. We

[3] In the pre-plan we also wrote that we would also add a dummy variable for being in the control group and interact Treated community with being in the control group. As there are no treated children in the control communities, however, the interaction terms drop out of the regression and no additional information is gained.

see that 14 percent of the children in control communities are starting lower secondary education. This number increases to 26 percent in the treated communities. The other main outcomes are more similar across samples and we see that children think the appropriate marriage rate is around 21 years, that over 40 percent of the children are coded as being religiously extreme, that children on average score above the low self-esteem cutoff (which is a Rosenberg index score of 15), and that around 20 percent of the children are involved in hazardous work. As the gender index is standardized, the numbers are close to zero, but we know that over 60 percent of the children in the control communities do not disagree that sending boys to secondary school is more important than sending girls, less than 15 percent think that girls should be allowed to study in secondary school if it is far away, and around 70 percent agree that men should have priority to remunerated work when it is rare.

With respect to the other outcome variables we see that 3 percent in the control communities did the Government test and say that they qualified for secondary education. Hence, there are fewer that report being qualified than have started secondary education. Most children in the control communities are both innumerate and illiterate despite having more than three years of schooling on average.

Regarding the control variables, we see that at baseline the children want large families, with almost 8 children on average. We also see that a large majority of the children live with male household heads and these household heads are almost 50 years old on average.

Table 1: Descriptive statistics

	All		Individual Treatment		Individual Control		Community Treatment		Community Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Treatment and strata variables										
Treatment individual	0.34	(0.48)	1.00	(0.00)	0.00	(0.00)	0.60	(0.49)	0.00	(0.00)
Treatment cluster	0.57	(0.50)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	0.00	(0.00)
Female	0.47	(0.50)	0.42	(0.49)	0.46	(0.50)	0.43	(0.50)	0.52	(0.50)
Main outcome variables										
Starting lower secondary	0.21	(0.40)	0.35	(0.48)	0.12	(0.32)	0.26	(0.44)	0.14	(0.34)
Appropriate marriage age	21.48	(4.04)	21.89	(3.78)	21.60	(4.28)	21.78	(3.98)	21.07	(4.08)
Religious extremism	0.45	(0.50)	0.46	(0.50)	0.47	(0.50)	0.47	(0.50)	0.44	(0.50)
Rosenberg index	19.96	(3.63)	20.02	(3.71)	20.10	(3.58)	20.05	(3.66)	19.84	(3.59)
Gender equality index	0.03	(1.01)	0.03	(1.02)	-0.00	(1.00)	0.02	(1.01)	0.04	(1.01)
Hazardous work	0.19	(0.40)	0.20	(0.40)	0.18	(0.39)	0.19	(0.40)	0.20	(0.40)
Other outcome variables										
Qualified for secondary	0.16	(0.37)	0.36	(0.48)	0.06	(0.24)	0.25	(0.44)	0.03	(0.18)
Numeracy index	1.08	(1.57)	1.56	(1.70)	0.87	(1.48)	1.28	(1.65)	0.82	(1.41)
Illiterate	0.80	(0.40)	0.68	(0.47)	0.86	(0.35)	0.75	(0.43)	0.86	(0.35)
Highest grade	4.46	(2.06)	4.64	(2.12)	4.32	(2.02)	4.53	(2.09)	4.36	(2.01)
Years of schooling	3.50	(2.80)	3.82	(2.80)	3.07	(2.76)	3.52	(2.81)	3.48	(2.80)
Social desirability	6.11	(1.60)	6.09	(1.61)	6.05	(1.62)	6.07	(1.62)	6.17	(1.59)
Baseline control variables										
Years of schooling	3.06	(2.65)	3.01	(2.66)	2.78	(2.63)	2.92	(2.65)	3.24	(2.64)
Desired fertility	7.71	(3.46)	7.73	(3.39)	7.71	(3.54)	7.72	(3.45)	7.69	(3.47)
Rosenberg index	20.05	(3.41)	20.10	(3.44)	19.84	(3.43)	20.00	(3.44)	20.12	(3.37)
Religious intolerance	2.58	(0.90)	2.59	(0.92)	2.55	(0.94)	2.57	(0.93)	2.60	(0.86)
Accept religious violence	1.93	(0.76)	1.98	(0.81)	1.90	(0.77)	1.95	(0.79)	1.91	(0.72)
Gender eq. work	0.39	(0.49)	0.35	(0.48)	0.40	(0.49)	0.37	(0.48)	0.41	(0.49)
Gender eq. abuse	0.34	(0.47)	0.33	(0.47)	0.34	(0.47)	0.34	(0.47)	0.34	(0.47)
Gender eq. school	0.45	(0.50)	0.44	(0.50)	0.44	(0.50)	0.44	(0.50)	0.47	(0.50)
Male hh head	0.85	(0.36)	0.87	(0.34)	0.86	(0.35)	0.87	(0.34)	0.82	(0.38)
HH asset index	-0.04	(1.05)	-0.06	(1.08)	-0.07	(1.11)	-0.06	(1.09)	-0.01	(0.99)
Age of hh head	49.14	(11.61)	49.20	(11.56)	48.81	(10.94)	49.05	(11.31)	49.26	(11.98)
N	2348		807		533		1340		1008	

Notes: The table shows means and standard deviations for our main pre-registered variables for the different samples. The samples for the individual treatment and control individuals are restricted to treated communities.

To test for baseline balance between treatment and control groups, we regress $Treatment_i$ on the baseline control variables, both individually and jointly, while controlling for the strata variables. The results, in Table 2, show that most variables are uncorrelated with treatment status. The only exceptions are that treated individuals seem more likely to accept religious violence and less likely to think that work should be prioritized to men if scarce at baseline. These correlations are only statistically significant at the 10 percent level and it is not surprising that some variables are statistically significantly different.

When we include all variables at the same time (in columns 2 and 4) the F-tests indicate that the variables do not jointly predict treatment status. As per our pre-plan, we therefore view the randomization as successful in creating balance and our main estimations are the ones with only the strata variables as controls. We also show that the results are similar when we add the full vector of controls and when choosing optimal controls. In Appendix Tables A.1 and A.2 we show descriptive statistics and balance tests for the full baseline samples as well, noting that also these samples are well balanced.

Table 2: Balance tests.

Outcome:	Individual Treatment		Community Treatment	
Controls added:	One-by-one	All	One-by-one	All
Year of schooling	0.0074	0.0072	-0.011	-0.010
	(0.0062)	(0.0063)	(0.011)	(0.010)
Desired fertility	0.000026	-0.00017	0.00070	0.00048
	(0.0045)	(0.0045)	(0.0049)	(0.0048)
Rosen index	0.0049	0.0055	-0.0025	-0.0021
	(0.0053)	(0.0056)	(0.0081)	(0.0073)
Religious intolerance	0.0075	0.0074	-0.0094	-0.011
	(0.016)	(0.017)	(0.020)	(0.019)
Accept religious violence	0.032*	0.028	0.020	0.018
	(0.019)	(0.020)	(0.025)	(0.023)
Gender eq. work	-0.053*	-0.058	-0.046	-0.038
	(0.032)	(0.036)	(0.035)	(0.032)
Gender eq. abuse	-0.0085	-0.0076	-0.0081	0.015
	(0.033)	(0.035)	(0.048)	(0.041)
Gender eq. school	0.0017	0.033	-0.026	-0.0069
	(0.031)	(0.034)	(0.035)	(0.032)
Male hh head	-0.0074	-0.017	0.085	0.098*
	(0.043)	(0.044)	(0.053)	(0.052)
HH asset index	-0.00071	-0.0026	-0.011	-0.012
	(0.014)	(0.014)	(0.020)	(0.019)
Age of hh head	0.00058	0.00079	-0.00039	-0.00075
	(0.0012)	(0.0013)	(0.0012)	(0.0012)
N	1340	1340	2348	2348
R-squared	NA	0.02	NA	0.02
Controls	Strata	Strata	Strata	Strata
P-value of F-test	NA	0.65	NA	0.58

Notes: All variables are measured at baseline. Each row in columns 1 and 3 shows coefficients from different regressions when we include the variables one by one. Columns 2 and 4 are from two regressions with all variables included. All regressions control for the strata variables. Robust SE in parentheses for the individual treatment regressions (columns 1 and 2) and clustered at the community level in the community treatment regressions (columns 3 and 4).

6 Effects on schooling and learning

In Table 3 we present our main outcome variable for schooling and learning outcomes: *Starting lower secondary education.* In column 1 the sample is restricted to the treated clusters and we use the individual level randomization. Since randomization was done within community and gender we always control for those variables in the individual level regressions (this is indicated in the table by "Strata" for control variables).

We see that the mean for the control group is 0.12, that is, 12 percent of the control kids in the treatment communities start lower secondary education. We see that there is an effect of individual treatment of 0.22. Hence, 34 percent of the treated kids start lower secondary and the effect is an 183 percent increase. As we pre-specified six main outcome variables it is also comforting to see that our p-value, which is lower than 0.0001, is lower than the critical p-value from a Bonferroni correction (0.008).[4] This result thereby also survives a correction for multiple hypotheses.

In columns 2 and 3 we run separate regressions by gender. This is seen in the "Gender" row at the bottom of the tables. We see that the effect is smaller for girls in absolute terms, but since the mean is also much lower the relative effect magnitude (238 percent) is much larger.

In column 4 we instead use the clustered random assignment and compare individuals randomly assigned to a treatment village to those randomly assigned to control villages. The strata variable in these regressions are region fixed effects as communities were randomly assigned within regions. Standard errors are here clustered at the community level as this is the level of the treatment variation. We see that individuals in treated villages are 12 percentage points more likely to start lower secondary education. This is almost a doubling from the 14 percent in the control villages.

[4]We pre-specified to use the Benjamini-Hochberg procedure for the multiple hypothesis testing, which is equivalent to a Bonferroni correction with only one statistically significant coefficient.

In column 5 we investigate spillover effects in the treated communities. We do this by restricting the sample to control kids in both types of communities and compare the kids randomly assigned to have a school in their village, that they were not randomly assigned to attend, to kids in control villages. As we are now comparing individuals we control for Female and community fixed effects as randomization was within these strata. There is no evidence of spillover effects.[5]

Table 3: Starting lower secondary education

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.22***	0.19***	0.24***		
	(0.021)	(0.029)	(0.030)		
Treatment cluster				0.12***	-0.020
				(0.031)	(0.026)
Female	-0.11***				-0.020
	(0.022)				(0.022)
Control mean	0.12	0.08	0.15	0.14	0.14
N	1340	579	761	2348	1541
R-squared	0.19	0.23	0.18	0.02	0.00
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

In Table 4 we turn to the treatment effects on learning outcomes. We see that there are clear and large effects at both the individual and community level. In this context, 86 percent of the control children cannot read and write and this is reduced by 17 percentage points (20 percent) for the treated children (panel a). Another way to interpret these numbers is that the literacy rate increased by 121 percent from 14 to 31 percentage points by being offered a place in the speed-school. The effects on numeracy are also impressive as we see in panel b. The control group children manage to solve 0.87 of the four tasks

[5] In Appendix Table A.3, we show the effects on three other measures of educational attainment and we see there are effects at both the individual and community level on being qualified for secondary school (panel a). There are also individual level effects on the highest grade attained and years of schooling but the effects at the community level are not statistically significant for these outcomes (panels b and c).

(addition, subtraction, multiplication, and division) on average. This increases to 1.52 for the treated children. This effect amounts to an increase of 0.44 of a standard deviation in the control group and 75 percent of the control group mean.

Table 4: Effects on learning outcomes.

(a) Illiterate

	(1)	(2)	(3)	(4)	(5)
Treatment individual	-0.17***	-0.15***	-0.18***		
	(0.021)	(0.028)	(0.031)		
Treatment cluster				-0.10***	0.0093
				(0.027)	(0.027)
Female	0.11***				0.050**
	(0.022)				(0.019)
Control mean	0.86	0.91	0.82	0.86	0.86
N	1340	579	761	2348	1541
R-squared	0.15	0.19	0.13	0.03	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

(b) Numeracy index

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.65***	0.48***	0.74***		
	(0.085)	(0.12)	(0.12)		
Treatment cluster				0.47***	0.034
				(0.11)	(0.12)
Female	-0.45***				-0.28***
	(0.085)				(0.073)
Control mean	0.87	0.70	1.01	0.82	0.82
N	1340	579	761	2348	1541
R-squared	0.15	0.16	0.17	0.03	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 3 and includes all individuals. The regression in column 5 follows equation 2 and the sample is restricted to control individuals. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Measuring effects on learning is important as attainment in general is sadly not very highly correlated with learning in most developing countries (Angrist et al., 2020). This is true also in our setting. Figure 1 shows the relationship between baseline schooling and learning outcomes for control and treated children in the treatment communities. Focusing first on the control children, we see that individuals without schooling are almost all illiterate which is not that surprising. More surprising is the fact that over 70 percent of control individuals that have been to school for 6 years are still illiterate and over half of those who have attended more than 6 years.[6] With respect to numeracy, we see that control children having attended school for 6 years can still on average only complete around one of the four numeracy tasks. The solid horizontal lines in the figures show the levels of illiteracy and numeracy for treated children and we note that being offered to attend speed school for two years causes the children to be better educated than children that have attended primary school for six years in the control group. The dashed horizontal lines show the average levels in the control group and these average levels are clearly lower than the average for the students having attended six years. We show in Appendix Figure A.4 that the results for starting secondary education follow the same pattern. We also see that there are clear treatment effects for all levels of previous education of the children and testing for heterogenous treatment effects by baseline years of schooling we find no statistically significant differences in the effects. Furthermore, there were no detectable heterogeneity in treatment effects when using the omnibus test proposed by Chernozhukov et al. (2018) in an honest random forest framework (Athey et al., 2019; Wager and Athey, 2018).

[6]Appendix Figures A.2 and A.3 show the distribution of years of education as well as in the three categories for the treated and control children in the treated communities. We can not reject that the distributions are equal (p-values 0.45 and 0.97 respectively).

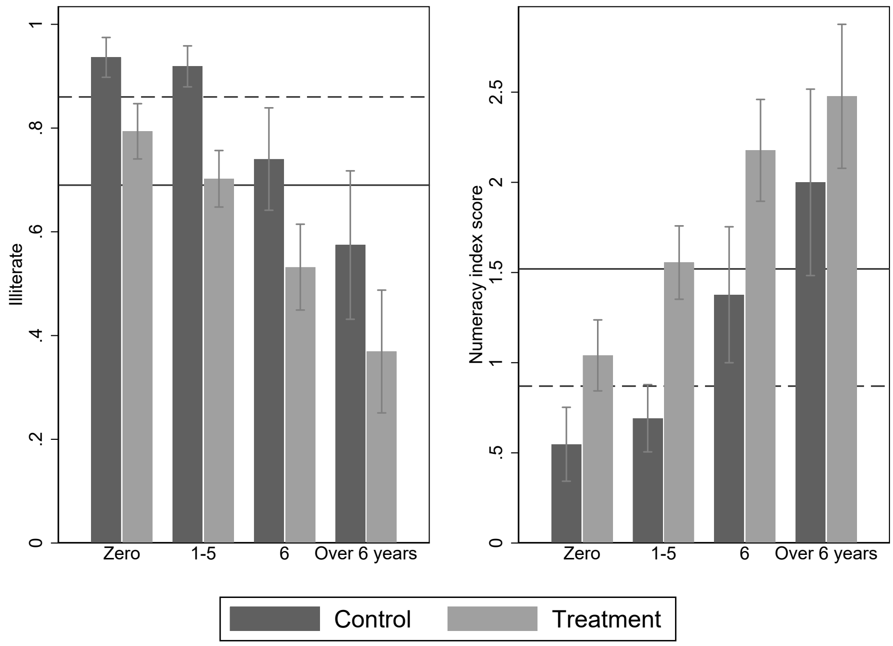

Figure 1: Relationship between years of schooling and learning outcomes for control and treated children.

Notes: The bars show the average levels of illiteracy and numeracy for control and treated children in the treatment communities by years of previous education with 95 percent confidence intervals. The solid horizontal lines show the average levels for children that were offered to attend speed school and the dashed solid lines the corresponding average for the control group.

In Figure 2 we show the distribution of the numeracy index separately for treated and control children in treatment communities. We see that treated children are more likely to be able to solve 1, 2, 3, and especially all 4 tasks. In Table 6, we show the individual level effects on the individual components of the numeracy index. We note that there are effects on all of them and that the effects on subtraction and multiplication are of over 100 percent. In Appendix Tables A.5 to A.8 we show the full tables for each of the numeracy variables and we note that the effects are often larger for boys than for girls, despite the means in the control groups already being higher for boys.

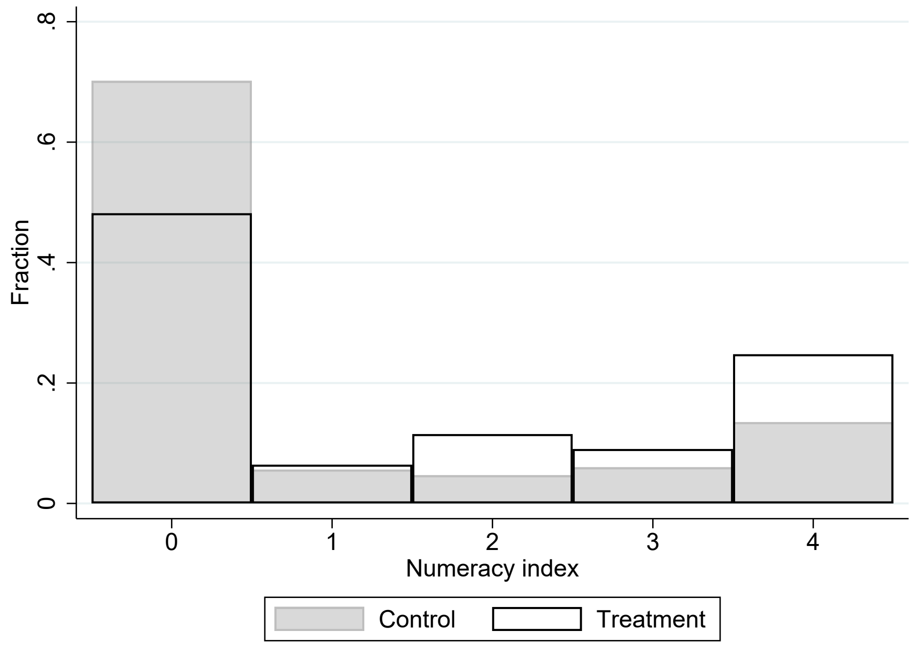

Figure 2: Distribution of the numeracy index for treated and control children.

Table 6: Individual level effects on different components of numeracy

	(1) Addition	(2) Subtraction	(3) Multiplication	(4) Division
Treatment individual	0.20***	0.19***	0.15***	0.12***
	(0.025)	(0.024)	(0.024)	(0.022)
Female	-0.13***	-0.11***	-0.12***	-0.085***
	(0.026)	(0.025)	(0.025)	(0.022)
Control mean	0.26	0.22	0.23	0.16
N	1340	1340	1340	1340
R-squared	0.14	0.14	0.12	0.12
Controls	Strata	Strata	Strata	Strata
Gender	Both	Both	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community. The regressions follow equation (1) and the sample is restricted to individuals in treated communities. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses.

6.1 Additional results and robustness

The results are similar both when adding the pre-specified vector of control variables and when using a double debiased LASSO approach to select optimal control variables. We show this in Appendix Tables A.9 and A.10 for starting lower secondary education but this conclusion holds for all variables investigated (see the populated pre-analysis plan).

In investigating attrition, we see in Table A.11 that there is a slight imbalance in attrition whereby treated individuals are less likely to attrit. This difference is only statistically significant at the 10 percent level. We also see that girls are less likely to attrit. When we add the set of extensive control variables there is no longer any statistically significant difference in attrition among treated and control children. The only variable that correlates with attrition when we add the full list of controls is years of schooling but the coefficient is very small. Taking the coefficient at face value it implies that children with one more year of schooling are 0.7 percent more likely to attrit. Conducting Lee bounds, only controlling for the strata variables, we note that the lower bounds of the treatment effect for starting lower secondary and for the learning outcomes are also statistically significant (see Appendix Table A.12).

In order to investigate social desirability bias, and more importantly, experimental demand effects we use a social desirability index (see Appendix Section A.2.1). The idea of the index is to see whether individuals are more likely to give saint-like answers if they are treated, which would then likely lead to self-reported measures being positively biased. Such worries are greater with reported attitudes than with outcomes from tests, but some of the educational outcomes are also self reported. As seen in Appendix Table A.18, there is no treatment effect on social desirability. We also note that there are no heterogeneous treatment effects based on the social desirability score (the p-value of the interaction term between social desirability and treatment is 0.5 in a regression on *Starting lower secondary education*). The worrisome pattern would be if the treatment effects were driven by children with a high propensity to disingenuously give socially desirable answers and vanished for those with a low such tendency.

6.2 Effects of attending the speed schools

Compliance to treatment assignment is of course not perfect and we did not expect that all children that were offered would attend. We do not have data on attendance that is recorded on a regular basis unfortunately but we do have administrative data on attendance from the schools during the first semester.[7] These data only separate between ever attended or not.

In Table 7 we see that 85 percent of the treated children attended the speed schools (column 1). The treatment effect on attending is similar for boys and girls (columns 2 and 3) and around half the children in treated clusters attend the schools (column 4). In column 5 we see no indications of spillover effects as control children in treated clusters are not more likely to have attended speed school than control children in control clusters. In Table A.13 we show the effects on attendance for the full sample, not only the endline sample, and the results are similar.

Table 7: Attended the speed school (SSA2) anytime during the first semester

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.85***	0.86***	0.84***		
	(0.013)	(0.020)	(0.018)		
Treatment cluster				0.51***	0.0014
				(0.012)	(0.0014)
Female	0.020				0.0011
	(0.015)				(0.0011)
Control mean	0.00	0.00	0.00	0.00	0.00
N	1340	579	761	2348	2161
R-squared	0.71	0.76	0.68	0.31	0.00
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

In Table 8 we present results where we predict attendance in speed school (SSA)

[7]We pre-registered to use a variable that equals one if the child answers yes on the question "In the past two years, have you been attending any school or education program?" and SSA2 on the question "What type of school or program did you attend?". This variable resulted in many children that we know actually attended SSA2 being coded as non-attendees, however, so we deviate from the plan to use the administrative records as our main measure of attendance. If we alternatively utilize survey-based measures of attendance, the effects amplify further as only 46 percent responded that they attended a speed school in the survey. We present these results in Appendix Table A.17.

with the randomization and use the predicted value for attending in the second stage to calculate the local average treatment effect (LATE) of attending on our outcomes. We see that the effects are larger, which is not surprising as the compliance rate is lower than 100 percent. The LATE for starting lower secondary education manifests as a 217 percent increase, while the LATE on illiteracy reflects a 23 percent decrease. Additionally, the LATE for numeracy skills is an impressive 89 percent of the control mean. When expressed in terms of standard deviations, the effect on numeracy equates to 52 percent of a standard deviation. This impact is considerable. For comparison, the median and 90th percentile learning effects from educational interventions in low and middle-income countries are 0.1 and 0.45 standard deviations respectively (Evans and Yuan, 2022).

In Appendix Table A.14 we also show the results by gender and we note that while the absolute levels of the effects are lower for girls, the relative effects as compared to the control children are larger.

Table 8: Instrumental variables estimates for the individual level treatment effects.

	(1) Starting LS	(2) Illiterate	(3) Numeracy
Attended SSA2	0.26***	-0.20***	0.77***
	(0.024)	(0.025)	(0.098)
Female	-0.11***	0.11***	-0.46***
	(0.022)	(0.022)	(0.084)
Control mean	0.12	0.86	0.87
N	1340	1340	1340
Controls	Strata	Strata	Strata
Gender	Both	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community. The estimates are based on instrumental variables regressions where the Attended SSA2 is first instrumented by individual treatment. The first stage is presented in Table A.13. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses.

7 Effects on other main outcomes

In addition to the results on educational outcomes we are also testing whether the program had any effects on other variables that are often hypothesized to be affected by education (such as outcomes related to gender equality, early marriage, and well-being) or that are particularly salient in our context (such as religious extremism or hazardous child work).

We see in Table 9 that there seem to be very limited effects on any of the other main outcome variables. The effects are not only mostly statistically insignificant, they are also small as compared to the means and standard deviations of the outcomes. The one instance where the effect is statistically significant is the community level effect on the appropriate marriage age in column 1 of panel b. While significant effects on 1/15 coefficients is likely to happen by chance alone it is also the case that this effect is relatively small. Taken at face value the coefficient implies that individuals in treated clusters think the appropriate marriage age is 0.7 years higher than the individuals in the control clusters. This effects amounts to around three percent of the control mean and is less than one fifth of the gender gap seen in column 1 of panel a. We conclude that being assigned to SSA improved educational attainment and learning but that it did not improve any of our other main outcomes.[8]

[8] There were no statistically significant effects for any variable when restricting the samples to either boys or girls, with the exception of there being a treatment effect on gender equality index for girls that is statistically significant at the 10 percent level (see Appendix Table A.16).

Table 9: Effects on main non-education outcomes.

(a) Individual level effects

	(1) Appropriate marriage age	(2) Religious extremism	(3) Rosenberg index	(4) Gender equality	(5) Hazardous work
Treatment individual	0.12	-0.012	-0.17	0.051	0.011
	(0.21)	(0.028)	(0.19)	(0.056)	(0.021)
Female	-3.58***	0.030	-0.029	0.30***	-0.16***
	(0.20)	(0.028)	(0.19)	(0.055)	(0.020)
Control mean	21.60	0.47	20.10	-0.00	0.18
N	1218	1329	1340	1322	1340
R-squared	0.24	0.06	0.14	0.07	0.09
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Both	Both	Both	Both

(b) Community level effects

	(1) Appropriate marriage age	(2) Religious extremism	(3) Rosenberg index	(4) Gender equality	(5) Hazardous work
Treatment cluster	0.70***	0.028	0.22	-0.016	-0.0017
	(0.25)	(0.028)	(0.30)	(0.051)	(0.022)
Control mean	21.07	0.44	19.84	0.04	0.20
N	2134	2320	2348	2316	2348
R-squared	0.01	0.01	0.01	0.00	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Both	Both	Both	Both

(c) Spillover effects

	(1) Appropriate marriage age	(2) Religious extremism	(3) Rosenberg index	(4) Gender equality	(5) Hazardous work
Treatment cluster	0.30	0.035	0.25	-0.020	-0.021
	(0.33)	(0.029)	(0.33)	(0.056)	(0.024)
Female	-3.86***	-0.031	-0.31*	0.28***	-0.13***
	(0.22)	(0.029)	(0.18)	(0.052)	(0.023)
Control mean	21.07	0.44	19.84	0.04	0.20
N	1396	1521	1541	1525	1541
R-squared	0.22	0.01	0.01	0.02	0.03
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Both	Both	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in Panel a, Region in panel b, and Region and gender in panel c. The regressions inpanel a follow equation (1) and the sample is restricted to individuals in treated communities. The regressions in panel b follow equation 3 and includes all individuals. The regressions in panel c follow equation 2 and the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in panel a clustered at the community level in panels b and c.

In Appendix section A.3 we see that there are no spillover effects on the siblings in terms of them being more likely to start or quit school, move, marry, or conduct hazardous work. Furthermore, there were no detectable heterogeneity in treatment effects on any of the other main variables either when using machine learning methods.

8 Conclusion

In the face of an ongoing learning crisis in low-income countries—where numerous students fail to master basic literacy and numeracy skills even after six years of schooling—we have evaluated the impact of an ambitious program, designed to prepare children for secondary school in just two years. Our findings reveal substantial improvements in both schooling and learning. Where only 12 percent of control group children in the treatment communities progressed to lower secondary education, this figure rose to 34 percent for those in the treatment group.

The program also significantly bolstered literacy rates, with an increase of 121 percent, and numeracy skills improved by 0.44 standard deviations. When we consider the less-than-perfect uptake of the intervention, we find that the program ranks among the top 10 percent of education interventions evaluated in middle and low-income countries (Evans and Yuan, 2022). Notably, no negative or positive spillover effects on schooling or learning were identified in the treated communities. In all, our results show that the learning crisis can be mitigated by supply of good schools and that neither demand for schooling nor even a lack of trained teachers seem to be binding constraints in our setting.

Beyond educational outcomes, we also examined whether the program influenced other factors often presumed to be affected by education—such as gender equality, early marriage, and well-being—as well as those specifically pertinent to our context, including religious extremism and hazardous child labor. However, the impact on these parameters was at most marginal. Most likely, effects on other aspects than learning need time to materialize and the hopes that education will be a short run silver bullet is likely unwarranted.

Given the low attendance rates and pervasive underperformance in formal schools, the need for catch-up programs targeting slightly older children is profound. The SSA2 program, costing the NGO 350 USD per child-year, is more expensive than the estimated average cost of 279 USD per child-year for primary education in Sub-Saharan Africa. However, considering the substantial learning gains yielded by SSA2, the additional cost appears justifiable.

We strongly encourage further research into the effects of speed schools in different contexts, as well as the exploration of cost-saving strategies. With further validation and optimization, programs like SSA2 could be scaled up to address the significant educational gap in low-income countries more effectively.

References

Aker, Jenny C and Christopher Ksoll, "Call me educated: Evidence from a mobile phone experiment in Niger," *Economics of Education Review*, 2019, *72*, 239–257.

Akyeampong, Kwame, Marcos Delprato, Ricardo Sabates, Zoe James, John Pryor, Jo Westbrook, Sarah Humphreys, and Haile Tsegay, "Speed School Programme in Ethiopia. Tracking the progress of Speed School students 2011-2017," 2018.

Angrist, Noam and Rachael Meager, "Implementation Matters: Generalizing Treatment Effects in Education," *Available at SSRN 4487496*, 2023.

_ , David K Evans, Deon Filmer, Rachel Glennerster, F Halsey Rogers, and Shwetlena Sabarwal, *How to Improve Education Outcomes Most Efficiently?: A Comparison of 150 Interventions Using the New Learning-adjusted Years of Schooling Metric* 2020.

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized random forests," *The Annals of Statistics*, 2019, *47* (2), 1148–1178.

Bagby, Emilie, Anca Dumitrescu, Cara Orfield, and Matt Sloan, "Niger IMAGINE long-term evaluation," *Washington, DC: Mathematica Policy Research*, 2016.

Baird, Sarah, J Aislinn Bohren, Craig McIntosh, and Berk Özler, "Optimal design of experiments in the presence of interference," *Review of Economics and Statistics*, 2018, *100* (5), 844–860.

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F Katz, Benjamin A Olken, and Anja Sautmann, "In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics," Technical Report, National Bureau of Economic Research 2020.

_ , Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying education: Evidence from two randomized experiments in India," *The quarterly journal of economics*, 2007, *122* (3), 1235–1264.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 2014, *81* (2), 608–650.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val, "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India," Technical Report, National Bureau of Economic Research 2018.

DHS, "Demographic and Health Survey 2012, Niger.," Technical Report 2012.

Di-Iorio-Daigle, Alexandre, "Impact des représentations du vieillissement et du sentiment de contrôle sur l'ajustement à la retraite dans une population de nouveaux retraités," 2019.

Diazgranados-Ferráns, Silvia, Jeongmin Lee, Chinedu Ohanyido, Kayla Hoyer, and Adane Miheretu, "The cost-effectiveness of an accelerated learning program on the literacy, numeracy and social-emotional learning outcomes of out-of-school children in northeast Nigeria: Evidence from a mixed methods randomized controlled trial," *Journal of Research on Educational Effectiveness*, 2022, *15* (4), 655–686.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Education, HIV, and early fertility: Experimental evidence from Kenya," *American Economic Review*, 2015, *105* (9), 2757–2797.

_ , _ , and _ , "The impact of free secondary education: Experimental evidence from Ghana," Technical Report, National Bureau of Economic Research 2021.

Eble, Alex, Chris Frost, Alpha Camara, Baboucarr Bouy, Momodou Bah, Maitri Sivaraman, Pei-Tseng Jenny Hsieh, Chitra Jayanty, Tony Brady, Piotr Gawron et al., "How much can we remedy very low learning levels in rural parts of low-income countries? Impact and generalizability of a multi-pronged para-teacher intervention from a cluster-randomized trial in the Gambia," *Journal of Development Economics*, 2021, *148*, 102539.

Evans, David K and Fei Yuan, "How big are effect sizes in international education studies?," *Educational Evaluation and Policy Analysis*, 2022, *44* (3), 532–540.

Fazzio, Ila, Alex Eble, Robin L Lumsdaine, Peter Boone, Baboucarr Bouy, Pei-Tseng Jenny Hsieh, Chitra Jayanty, Simon Johnson, and Ana Filipa Silva, "Large learning gains in pockets of extreme poverty: Experimental evidence from Guinea Bissau," *Journal of Public Economics*, 2021, *199*, 104385.

Glewwe, Paul and Karthik Muralidharan, "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 653–743.

INS, "RAPPORT Smart 2022: Enquête Nationale nutritionnelle et de mortalité rérospective au Niger.," *Institut National de la Statistique*, 2022.

_ , "Enquête Nationale sur la Fécondité et la Mortalité des Enfants de Moins de Cinq Ans 2021.," *Institut National de la Statistique*, 2023.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, "The challenge of education and learning in the developing world," *Science*, 2013, *340* (6130), 297–300.

Moussa, Issaka, "Stratégie de scolarisation accelerее 2eme formule (SSA2) et qualite de l'education au Niger.," *Doctoral Thesis. Ecole Normale Superieure, l'Universite Abdou Moumouni, Niamey*, Forthcoming.

Ozer, Simon and Preben Bertelsen, "Capturing violent radicalization: Developing and validating scales measuring central aspects of radicalization," *Scandinavian Journal of Psychology*, 2018, *59* (6), 653–660.

Pritchett, Lant, *The rebirth of education: Schooling ain't learning*, CGD Books, 2013.

UNESCO, "Global Education Monitoring Report.," Technical Report, United Nations Educational, Scientific and Cultural Organization 2021.

Vallieres, Evelyne F and Robert J Vallerand, "Traduction et validation canadienne-française de l'échelle de l'estime de soi de Rosenberg," *International journal of psychology*, 1990, *25* (2), 305–316.

Wager, Stefan and Susan Athey, "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 2018, *113* (523), 1228–1242.

World Bank, *World development report 2018: Learning to realize education's promise*, The World Bank, 2018.

_ , "World Bank Data - Niger. https://data.worldbank.org/country/NE," 2023.

Appendix:

A.1 Coding of variables

A.1.0.1 Schooling and learning outcomes

The overarching program goal of SSA2 is for out-of-school children to transfer back into lower secondary level of the traditional schooling system in Niger. Lower secondary school is divided into academic and vocational tracks (CEG/CET and CFM).

Our main measure for the educational outcomes is *Starting lower-secondary education*. This variable is coded as zero for everyone answering No on the question “Are you enrolled in school or in an education program this fall?”. It is coded as 1 if they answer yes and then answer CEG/CET or CFM on the follow up question “What type of education program is that?”. We also replace the variable to be equal to one if the respondent answers lower secondary school on the question on the highest grade level achieved.

We also analyze a number of secondary variables for education and learning. We code *Qualified for secondary* equal to 1 if the child answers yes to the question “Did you do the Government test to qualify for lower secondary this summer?” and answer CEG/CET or CFM on the question “What was the result of the Government test?”. Other secondary outcomes are years of schooling, highest grade level achieved, and results from reading, writing and math tests. We code the following:

Years of schooling is coded as the numerical answers to the question, “For how many years did you go to school, not counting pre-school?”. Missing values are replaced by zero for individuals answering No to the screening question “Have you ever been to school?”. *Highest grade level* ranges between 1-7 where 7 refers to all levels above 6.

We have several variables intended to measure learning. First we code a variable, *Illiterate*, to be equal to one if children answer no on the question “Can you read and write French?". If they answer yes we code them as zero if they manage to write their own name on a piece of paper we provide them and if they can read the sentence "Fatima likes going to school, but she also likes to help her father". We also create a *Numeracy index* ranging from 0 to 4 where one point is given for being able to solve each of the problems we give them on addition (77+11= ?), subtraction (49-12= ?), division (28/4= ?), and multiplication (4x5= ?).

To investigate compliance we code a variable, *Attended the speed school*, which is based on administrative data on attendance from the schools during the first semester.[9]

A.1.0.2 Other main outcomes

In this section we describe the coding of our main variables for each topic. For each topic we also registered a number of secondary analyses. We only present the results of these secondary analyses if there are effects on the main outcomes. Interested readers can see the coding in the pre-analysis plan (added to the Appendix) and results can be seen in our populated pre-analysis plan[10]

For marriage and fertility our main measure is *Appropriate marriage age* which is equal to the answer on the question "What do you think would be a good age for you to marry?". Non-numerical answers are coded as missing.

Religious extremism is a function of (1) religious intolerance and (2) acceptance of violent and/or illegal means (Ozer and Bertelsen, 2018). We therefore create a composite variable based on these two aspects. With regards to acceptance for violence, our proxy is a continuous variable for *Accept religious violence*: "In certain cases, it is acceptable

[9]We pre-registered to use a variable that equals one if the child answers yes on the question "In the past two years, have you been attending any school or education program?" and SSA2 on the question "What type of school or program did you attend?". This variable resulted in many children that we know actually attended SSA2 being coded as non-attendees, however, so we deviate from the plan to use the administrative records as our main measure of attendance. We present the results using the pre-specified survey measure as well.

[10]Available at https://andreaskotsadam.wordpress.com/wp-content/uploads/2024/06/populated-pap-ssa-niger.pdf

to use violence to defend religious principles", which is coded on a scale from 0 (Strongly disagree) to 10 (Strongly agree). With regards to *Religious intolerance*, we use a continuous variable based on: "If my friend abandons the faith, I will avoid him/her", also coded on a range from 0 (Strongly disagree) to 10 (Strongly agree). The variable *Religious extremism* is then a dummy variable equal to 1 if each indicator gets at least a score of 5. To make a more precise proxy for potential support to violent groups we recode to zero respondents who on the question "Why do you think some young people these days chose to join an armed group" select response categories a; "to gain personal power", b; "to enrich themselves" or j; "they are bandits". Similarly, we code to zero those who, to the question "What should government do to prevent young people from joining armed, violent groups" respond a) "prohibit radical preaching".

Our main outcome to measure self-esteem and wellbeing is the *Rosenberg index*. In particular, we use a validated French translation of the Rosenberg self-esteem scale (Vallieres and Vallerand, 1990) which contains the following 10 survey questions:

- I am generally satisfied with myself.
- Sometimes, I think I'm no good.
- I think a have several good qualities
- I'm able to do the same things as other people
- I don't have much to be proud of
- Sometimes, I certainly feel useless
- I feel that I'm as valuable as anyone else
- I wish I had more self respect
- I generally think that I'm a failure.
- I have a positive attitude to myself

The answer alternatives are: 0 Strongly agree, 1 Agree, 2 Disagree, 3 Strongly disagree. Items 2, 5, 6, 8, 9 are reverse scored. We therefore give "Strongly Disagree" 0 points, "Disagree" 1 point, "Agree" 2 points, and "Strongly Agree" 3 points on these items. We then sum the scores for all ten items and keep scores on a continuous scale. Higher scores indicate higher self-esteem. A score below 15 (out of the total 30) is usually interpreted as low self esteem.

For hazardous child labor and time use our main variable is a dummy variable, *Hazardous child labor*, which is equal to 1 if the child answers that they during the past two years have worked in at least one of the following types of work places:

- A household of someone who were not a relative
- A mine
- A quarry
- The streets in a city
- On a farm handling chemicals
- On a farm where one could sense the smell of chemicals
- Carrying loads that one felt were painfully heavy

Our main measure of gender attitudes is an index. We have a section in the survey with attitude questions after the prompt: "I will read out some different statements to you, and I would like for you to tell me if you agree or not. You should answer exactly what you'd like. There are no right or wrong answers." The answer categories are from 0 to 10 where 0 is strongly disagree and 10 is strongly agree. Our main outcome is *Gender equality index* and is based on dummy coding of the following variables:

- It is more important to send boys to secondary school
- Girls should be allowed to study in higher secondary school even if it is far away
- When remunerated work is rare, men should have priority to the jobs

The dummy codings follow our general rule for such codings where we split the variables so that we get as equal sized groups as possible but we code the dummies so that 1 is always the more gender equal answers.

A.1.0.3 Main independent variables

All main independent variables are from the baseline surveys. We always have the strata variables (female and community fixed effects for the individual analysis and region fixed effects for the community analysis) as independent variables.

The other main independent variables are continuous numerical values of:

- Years of schooling: Already defined.
- Desired fertility: Already defined.
- Rosenberg index of self-esteem: Already defined.
- HH respondent age: Numerical value of the age of the household respondent.
- Asset index: Based on the principal component analysis of ownership of 12 items.[11]

The other main independent variables are also dummy variables for:

- Religious intolerance: "If my friend abandons the faith, I will avoid him/her."
- Accept religious violence: "It is acceptable to use violence to defend religious principles."
- Gender Equality school: "It is more important to send boys to secondary school"
- Gender Equality work: "When remunerated work is rare, men should have priority to the jobs"
- Gender Equality abuse: "It is justifiable for a man to beat a woman if she: burns the food"
- HH respondent male: Equal to 1 if the household respondent is male.

[11]Radio, Fridge, Cellphone, Bicycle, Moped, Motorbike, Car, Electricity, Solar panel, Generator, Push cart, Cistern or water reservoir. The asset index is standardized by subtracting the mean of the control communities and dividing by the standard deviation of the control communities.

A.2 Figures and tables referred to in the text

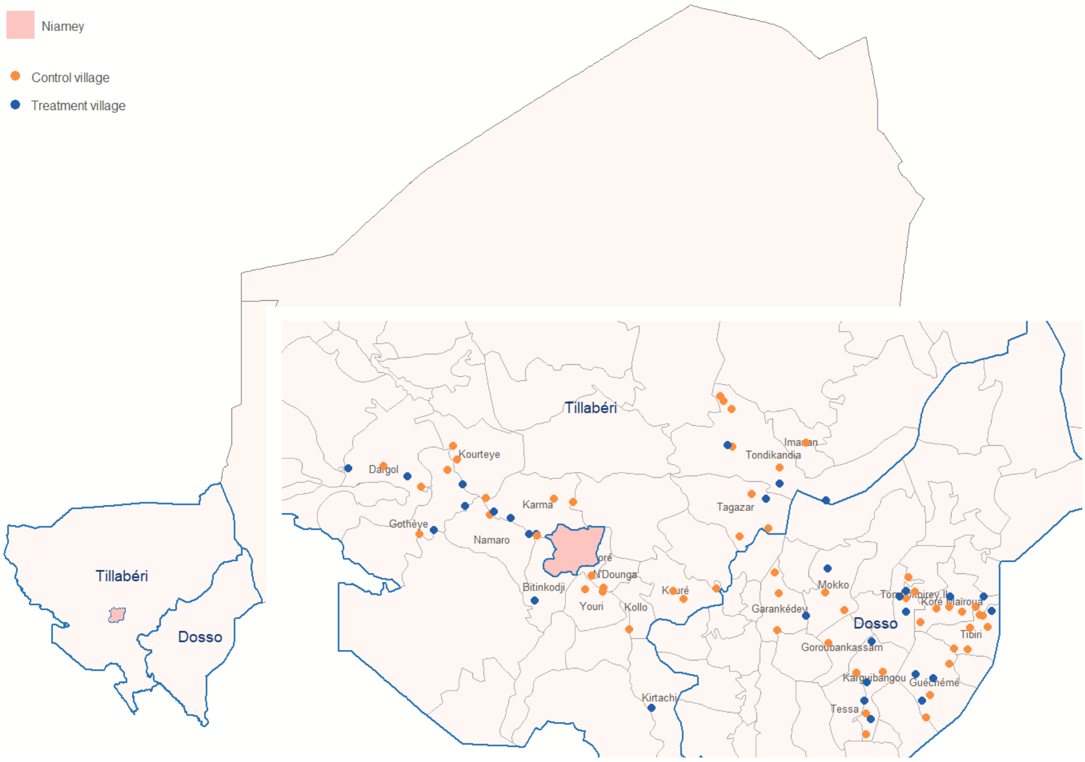

Figure A.1: Map of treated and control villages.

Table A.1: Descriptive statistics for the full baseline sample

	All		Individual Treatment		Individual Control		Community Treatment		Community Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Treatment and strata variables										
Treatment individual	0.33	(0.47)	1.00	(0.00)	0.00	(0.00)	0.59	(0.49)	0.00	(0.00)
Treatment cluster	0.56	(0.50)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	0.00	(0.00)
Female	0.45	(0.50)	0.40	(0.49)	0.44	(0.50)	0.42	(0.49)	0.50	(0.50)
Main outcome variables										
Starting lower secondary	0.21	(0.40)	0.35	(0.48)	0.12	(0.32)	0.26	(0.44)	0.14	(0.34)
Appropriate marriage age	21.48	(4.04)	21.89	(3.78)	21.60	(4.28)	21.78	(3.98)	21.07	(4.08)
Religious extremism	0.45	(0.50)	0.46	(0.50)	0.47	(0.50)	0.47	(0.50)	0.44	(0.50)
Rosenberg index	19.96	(3.63)	20.02	(3.71)	20.10	(3.58)	20.05	(3.66)	19.84	(3.59)
Gender equality index	0.03	(1.01)	0.03	(1.02)	-0.00	(1.00)	0.02	(1.01)	0.04	(1.01)
Hazardous work	0.19	(0.40)	0.20	(0.40)	0.18	(0.39)	0.19	(0.40)	0.20	(0.40)
Other outcome variables										
Qualified for secondary	0.16	(0.37)	0.36	(0.48)	0.06	(0.24)	0.25	(0.44)	0.03	(0.18)
Numeracy index	1.08	(1.57)	1.56	(1.70)	0.87	(1.48)	1.28	(1.65)	0.82	(1.41)
Illiterate	0.80	(0.40)	0.68	(0.47)	0.86	(0.35)	0.75	(0.43)	0.86	(0.35)
Highest grade	4.46	(2.06)	4.64	(2.12)	4.32	(2.02)	4.53	(2.09)	4.36	(2.01)
Years of schooling	3.50	(2.80)	3.82	(2.80)	3.07	(2.76)	3.52	(2.81)	3.48	(2.80)
Social desirability	6.11	(1.60)	6.09	(1.61)	6.05	(1.62)	6.07	(1.62)	6.17	(1.59)
Baseline control variables										
Years of schooling	3.11	(2.67)	3.02	(2.68)	2.91	(2.65)	2.98	(2.67)	3.28	(2.67)
Desired fertility	7.71	(3.46)	7.73	(3.39)	7.71	(3.54)	7.72	(3.45)	7.69	(3.47)
Rosenberg index	20.01	(3.41)	20.05	(3.40)	19.90	(3.48)	19.99	(3.43)	20.02	(3.39)
Religious intolerance	2.58	(0.90)	2.57	(0.90)	2.56	(0.95)	2.56	(0.92)	2.60	(0.87)
Accept religious violence	1.93	(0.75)	1.98	(0.79)	1.88	(0.76)	1.94	(0.78)	1.91	(0.72)
Gender eq. work	0.38	(0.49)	0.35	(0.48)	0.40	(0.49)	0.37	(0.48)	0.40	(0.49)
Gender eq. abuse	0.33	(0.47)	0.32	(0.47)	0.34	(0.48)	0.33	(0.47)	0.33	(0.47)
Gender eq. school	0.45	(0.50)	0.44	(0.50)	0.45	(0.50)	0.45	(0.50)	0.46	(0.50)
Male hh head	0.85	(0.36)	0.86	(0.35)	0.86	(0.34)	0.86	(0.35)	0.83	(0.38)
HH asset index	-0.03	(1.04)	-0.05	(1.05)	-0.06	(1.10)	-0.05	(1.07)	0.00	(1.00)
Age of hh head	49.44	(11.74)	49.48	(11.79)	48.97	(11.00)	49.27	(11.48)	49.65	(12.05)
N	3216		1055		731		1786		1430	

Notes: The table shows means and standard deviations for our main pre-registered variables for the different samples. The samples for the individual treatment and control individuals are restricted to treated communities.

Table A.2: Balance tests for the full baseline sample

Outcome:	Individual Treatment		Community Treatment	
Controls added:	One-by-one	All	One-by-one	All
Year of schooling	0.0025	0.0033	-0.010	-0.0093
	(0.0054)	(0.0055)	(0.0099)	(0.0097)
Desired fertility	-0.00056	-0.00057	0.00068	0.00024
	(0.0044)	(0.0044)	(0.0049)	(0.0048)
Rosenberg index	0.0034	0.0047	-0.00051	-0.000014
	(0.0047)	(0.0049)	(0.0074)	(0.0067)
Religious intolerance	0.0020	0.0023	-0.012	-0.011
	(0.015)	(0.015)	(0.019)	(0.018)
Accept religious violence	0.041**	0.038**	0.014	0.015
	(0.017)	(0.018)	(0.025)	(0.022)
Gender eq. work	-0.045	-0.042	-0.026	-0.021
	(0.028)	(0.031)	(0.033)	(0.031)
Gender eq. abuse	-0.029	-0.027	-0.0032	0.010
	(0.029)	(0.031)	(0.047)	(0.039)
Gender eq. school	-0.0070	0.019	-0.014	-0.0066
	(0.027)	(0.030)	(0.034)	(0.030)
Male hh head	-0.025	-0.038	0.064	0.077
	(0.037)	(0.038)	(0.052)	(0.052)
HH asset index	0.00052	-0.00061	-0.011	-0.012
	(0.012)	(0.013)	(0.019)	(0.018)
Age of hh head	0.00084	0.0012	-0.00066	-0.00091
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
N	1786	1786	3216	3216
R-squared	NA	0.01	NA	0.01
Controls	Strata	Strata	Strata	Strata
P-value of F-test	NA	0.43	NA	0.75

Notes: All variables are measured at baseline. Each row in columns 1 and 3 shows coefficients from different regressions when we include the variables one by one. Columns 2 and 4 are from two regressions with all variables included. All regressions control for the strata variables. Robust SE in parentheses for the individual treatment regressions (columns 1 and 2) and clustered at the community level in the community treatment regressions (columns 3 and 4).

Table A.3: Effects on other measures of educational attainment.

(a) Qualified for secondary education

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.30***	0.26***	0.31***		
	(0.023)	(0.035)	(0.031)		
Treatment cluster				0.22***	0.025*
				(0.025)	(0.014)
Female	-0.044*				-0.020*
	(0.026)				(0.011)
Control mean	0.06	0.05	0.07	0.03	0.03
N	1010	402	608	1778	1134
R-squared	0.19	0.24	0.20	0.09	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

(b) Highest grade attained

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.31**	0.24	0.39**		
	(0.12)	(0.19)	(0.16)		
Treatment cluster				0.16	-0.065
				(0.21)	(0.23)
Female	-0.36***				-0.26*
	(0.13)				(0.13)
Control mean	4.32	4.13	4.45	4.36	4.36
N	1010	402	608	1778	1134
R-squared	0.19	0.33	0.15	0.00	0.00
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

(c) Years of schooling

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.68***	0.69***	0.66***		
	(0.14)	(0.20)	(0.19)		
Treatment cluster				0.041	-0.45
				(0.28)	(0.28)
Female	-0.74***				-0.59***
	(0.14)				(0.17)
Control mean	3.07	2.62	3.45	3.48	3.48
N	1340	579	761	2348	1541
R-squared	0.23	0.30	0.20	0.01	0.02
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.5: Addition

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.20***	0.16***	0.22***		
	(0.025)	(0.037)	(0.035)		
Treatment cluster				0.12***	-0.0090
				(0.031)	(0.035)
Female	-0.13***				-0.092***
	(0.026)				(0.024)
Control mean	0.26	0.21	0.30	0.26	0.26
N	1340	579	761	2348	1541
R-squared	0.14	0.14	0.16	0.02	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.6: Subtraction

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.19***	0.16***	0.20***		
	(0.024)	(0.035)	(0.034)		
Treatment cluster				0.14***	0.016
				(0.030)	(0.031)
Female	-0.11***				-0.076***
	(0.025)				(0.021)
Control mean	0.22	0.17	0.27	0.20	0.20
N	1340	579	761	2348	1541
R-squared	0.14	0.14	0.17	0.03	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.7: Multiplication

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.15***	0.071**	0.19***		
	(0.024)	(0.036)	(0.034)		
Treatment cluster				0.12***	0.019
				(0.028)	(0.033)
Female	-0.12***				-0.063***
	(0.025)				(0.019)
Control mean	0.23	0.20	0.26	0.21	0.21
N	1340	579	761	2348	1541
R-squared	0.12	0.12	0.15	0.02	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.8: Division

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.12***	0.090***	0.14***		
	(0.022)	(0.030)	(0.031)		
Treatment cluster				0.085***	0.0081
				(0.026)	(0.027)
Female	-0.085***				-0.045**
	(0.022)				(0.018)
Control mean	0.16	0.12	0.19	0.15	0.15
N	1340	579	761	2348	1541
R-squared	0.12	0.13	0.13	0.02	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

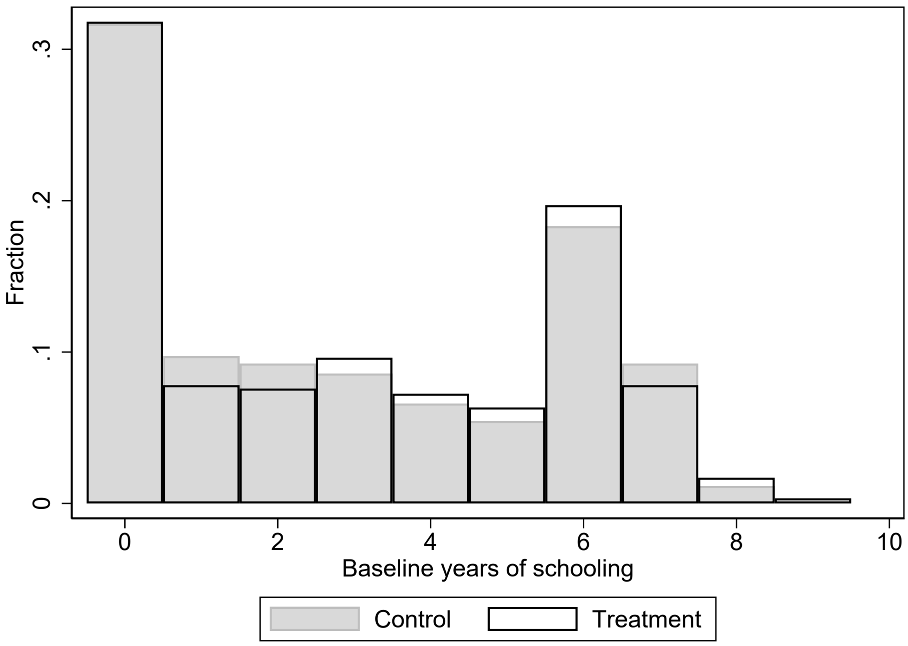

Figure A.2: Baseline years of schooling in three categories for control and treated children.

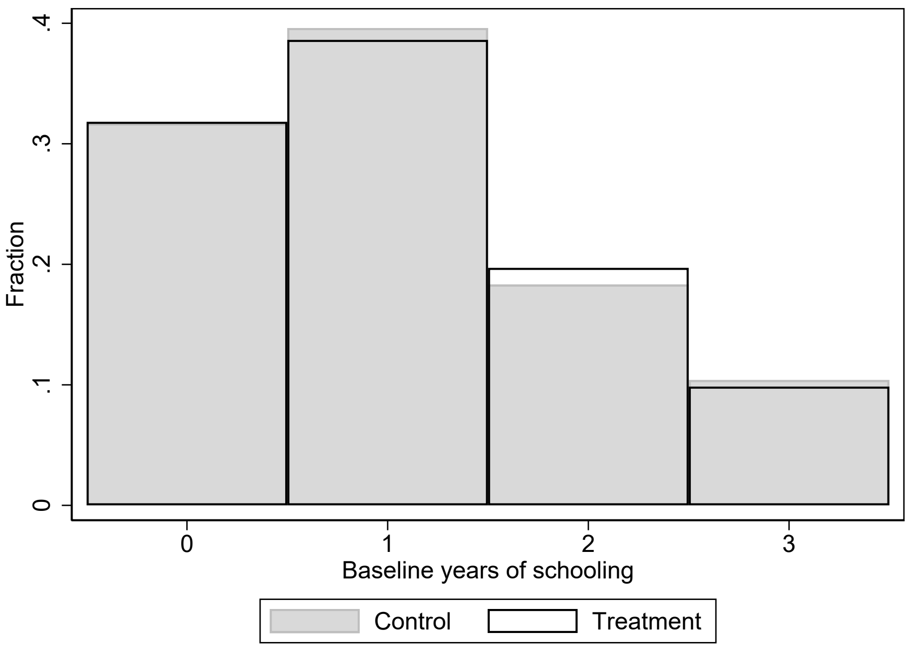

Figure A.3: Baseline years of schooling in three categories for control and treated children.

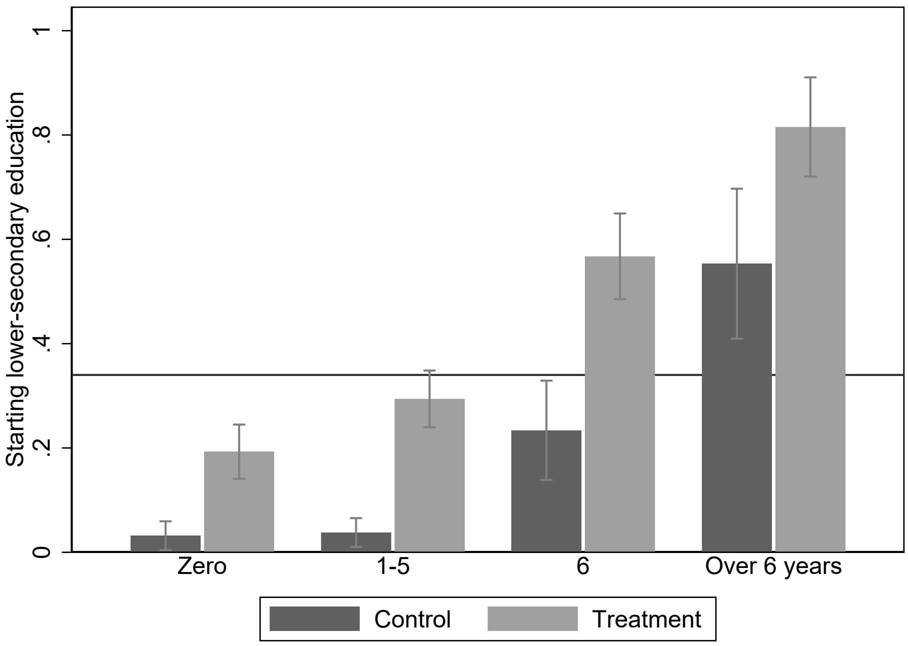

Figure A.4: Relationship between years of schooling and starting secondary education for control and treated children.

Notes: The bars show the average levels of secondary school attendance for control and treated children in the treatment communities by years of previous education with 95 percent confidence intervals. The solid horizontal line show the average levels for children that were offered to attend speed school and the dashed solid line shows the corresponding average for the control group.

Table A.9: Starting lower secondary education. Results with controls

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.21***	0.19***	0.23***		
	(0.020)	(0.027)	(0.029)		
Treatment cluster				0.14***	0.0051
				(0.024)	(0.021)
Female	-0.068***				0.0051
	(0.023)				(0.019)
Control mean	0.12	0.08	0.15	0.14	0.14
N	1340	579	761	2348	1541
R-squared	0.27	0.33	0.26	0.15	0.15
Controls	Extensive	Extensive	Extensive	Extensive	Extensive
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for the full set of controls described in Section A.1.0.3 and for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.10: Starting lower secondary education. Results with optimal controls

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.21***	0.19***	0.22***		
	(0.020)	(0.026)	(0.028)		
Treatment cluster				0.14***	0.0012
				(0.025)	(0.019)
Female	-0.070***				0.0018
	(0.021)				(0.019)
Control mean	0.12	0.08	0.15	0.14	0.14
N	1340	579	761	2348	1541
Controls	Lasso	Lasso	Lasso	Lasso	Lasso
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for the optimal controls as chosen by a double debiased LASSO procedure and for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.11: Attrition

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment individual	-0.036*	-0.038	-0.039			-0.0097
	(0.021)	(0.030)	(0.028)			(0.013)
Treatment cluster				-0.044*	-0.027	
				(0.025)	(0.025)	
Female	-0.046**				-0.058**	-0.00067
	(0.020)				(0.024)	(0.013)
Years of schooling						0.0071**
						(0.0029)
Desired fertility						0.0011
						(0.00081)
Rosenberg index						0.0015
						(0.0024)
Religious intolerance						0.0069
						(0.0072)
Accept religious violence						0.0021
						(0.0082)
Gender eq. work						0.023
						(0.015)
Gender eq. abuse						0.013
						(0.014)
Gender eq. school						-0.017
						(0.015)
Male hh head						-0.031
						(0.020)
HH asset index						0.0068
						(0.0053)
Age of hh head						0.00022
						(0.00060)
Control mean	0.27	0.24	0.29	0.29	0.29	0.27
N	1786	744	1042	3216	2161	1786
R-squared	0.05	0.08	0.05	0.01	0.01	0.66
Controls	Strata	Strata	Strata	Strata	Strata	Extensive
Gender	Both	Girls	Boys	Both	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. In column 6 we follow equation (1) but also add the full list of controls. P-values are ≤ 0.01***, ≤ 0.05**, and ≤ 0.1*. Robust SE in parentheses in columns 1-3 and column 6 clustered at the community level in columns 4 and 5.

Table A.12: Lee bounds of the individual treatment effects

	Coefficient	Std. err.	P>|z|
Starting lower secondary education			
Bounds			
lower	0.21	0.06	0.001
upper	0.24	0.07	0.001
Numeracy index			
Bounds			
lower	0.59	0.24	0.014
upper	0.72	0.27	0.009
Illiterate			
Bounds			
lower	-0.19	0.07	0.005
upper	-0.15	0.06	0.018

Notes: Upper and lower bounds for the main education and learning outcomes using Lee bounds. Controlling for gender and community fixed effects.

Table A.13: Attended SSA. Using the full sample

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.81***	0.81***	0.81***		
	(0.012)	(0.019)	(0.016)		
Treatment cluster				0.48***	0.0014
				(0.010)	(0.0014)
Female	0.0056				0.0011
	(0.015)				(0.0011)
Control mean	0.00	0.00	0.00	0.00	0.00
N	1786	744	1042	3216	2161
R-squared	0.65	0.68	0.64	0.29	0.00
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.14: Instrumental variables estimates for the individual level treatment effects.

(a) Starting lower secondary education

	(1)	(2)	(3)
Attended SSA2	0.26***	0.22***	0.29***
	(0.024)	(0.033)	(0.035)
Female	-0.11***		
	(0.022)		
Control mean	0.12	0.08	0.15
N	1340	579	761
Controls	Strata	Strata	Strata
Gender	Both	Girls	Boys

(b) Illiterate

	(1)	(2)	(3)
Attended SSA2	-0.20***	-0.17***	-0.22***
	(0.025)	(0.032)	(0.037)
Female	0.11***		
	(0.022)		
Control mean	0.86	0.91	0.82
N	1340	579	761
Controls	Strata	Strata	Strata
Gender	Both	Girls	Boys

(c) Numeracy index

	(1)	(2)	(3)
Attended SSA2	0.77***	0.56***	0.89***
	(0.098)	(0.13)	(0.14)
Female	-0.46***		
	(0.084)		
Control mean	0.87	0.70	1.01
N	1340	579	761
Controls	Strata	Strata	Strata
Gender	Both	Girls	Boys

Notes: All regressions control for strata fixed effects, which are gender and community. The estimates are based on instrumental variables regressions where the Attended SSA2 is first instrumented by individual treatment. The first stage is presented in Table A.13. The sample consists of only girls in column 2, and of only boys in column 3. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses.

Table A.16: Gender equality index

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.051	0.15*	-0.014		
	(0.056)	(0.086)	(0.074)		
Treatment cluster				-0.016	-0.020
				(0.051)	(0.056)
Female	0.30***				0.28***
	(0.055)				(0.052)
Control mean	-0.00	0.10	-0.09	0.04	0.04
N	1322	570	752	2316	1525
R-squared	0.07	0.10	0.06	0.00	0.02
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.17: Instrumental variables estimates for the individual level treatment effects using pre-registered measure of attendance.

	(1)	(2)	(3)
	Starting LS	Illiterate	Numeracy
Attended SSA2	0.58***	-0.45***	1.70***
	(0.056)	(0.055)	(0.22)
Female	-0.077***	0.087***	-0.37***
	(0.023)	(0.022)	(0.086)
Control mean	0.12	0.86	0.87
N	1340	1340	1340
Controls	Strata	Strata	Strata
Gender	Both	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community. The estimates are based on instrumental variables regressions where the Attended SSA2 is first instrumented by individual treatment. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses.

A.2.1 Social desirability

We use a *Social desirability index* to investigate whether experimenter demand effects are upward biasing the estimated program impacts. The index is based on responses on a Marlowe-Crowne module which measures a person's general tendency to give socially desirable answers. We use a validated French translation (Di-Iorio-Daigle, 2019). The index is based on the following questions where individuals can answer agree or disagree:

1. It is sometimes hard for me to go on with my work if I am not encouraged.
2. I sometimes feel resentful when I don't get my way.
3. On a few occasions, I have given up doing something because I thought too little of my ability.
4. There have been times when I felt like rebelling against people in positions of authority even though I knew they were right.
5. No matter who I'm talking to, I'm always a good listener.
6. There have been occasions when I took advantage of someone.
7. I am always willing to admit it when I make a mistake.
8. I sometimes try to get even rather than forgive and forget.
9. I am always courteous, even to people who are disagreeable.
10. I have never been irked when people expressed ideas very different from my own.
11. There have been times when I was quite jealous of the good fortune of others.
12. I am sometimes irritated by people who ask favours of me.
13. I have never deliberately said something that hurt someone's feelings.

Table A.18: Social desirability

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.028	-0.090	0.13		
	(0.088)	(0.13)	(0.12)		
Treatment cluster				-0.096	-0.12
				(0.10)	(0.11)
Female	0.067				0.15
	(0.086)				(0.093)
Control mean	6.05	6.16	5.95	6.17	6.17
N	1340	579	761	2348	1541
R-squared	0.09	0.12	0.09	0.00	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

A.3 Spillover effects on siblings

In investigating spillover effects we additionally investigate if there are treatment effects on the following variables that were asked about what siblings (same mother and/or same father) 7-18 years old were doing in the past two years:

"We would now like for you to think about what these siblings were doing in the past two years."

- Did any of your siblings start school in the fall of 2019 (last year)?
- Did any of your siblings quit school in the past two years?
- Did any of your siblings under the age of 18 leave the village to earn money in the past two years?
- Did any of your siblings under the age of 18 leave the country during the past two years?
- Did any of your siblings under the age of 18 marry during the past two years?
- Did any of your siblings under the age of 18 go to work in a mine or a quarry during the past two years?

The answers are dummy coded.

Table A.19: Sibling started school

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.012	-0.038	0.042		
	(0.029)	(0.045)	(0.040)		
Treatment cluster				-0.017	-0.023
				(0.036)	(0.042)
Female	-0.0091				0.0015
	(0.029)				(0.029)
Control mean	0.55	0.58	0.53	0.58	0.58
N	1153	498	655	2009	1306
R-squared	0.08	0.11	0.10	0.00	0.00
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.20: Sibling quitted school

	(1)	(2)	(3)	(4)	(5)
Treatment individual	-0.021	0.0026	-0.046		
	(0.027)	(0.042)	(0.035)		
Treatment cluster				-0.023	-0.0088
				(0.027)	(0.032)
Female	0.0030				-0.027
	(0.027)				(0.027)
Control mean	0.31	0.29	0.32	0.32	0.32
N	1153	498	655	2009	1306
R-squared	0.08	0.09	0.13	0.01	0.02
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.21: Sibling moved to earn money

	(1)	(2)	(3)	(4)	(5)
Treatment individual	-0.015	0.0067	-0.034		
	(0.024)	(0.034)	(0.034)		
Treatment cluster				0.00022	0.0045
				(0.023)	(0.027)
Female	-0.057**				-0.058**
	(0.024)				(0.025)
Control mean	0.20	0.17	0.23	0.20	0.20
N	1153	498	655	2009	1306
R-squared	0.05	0.11	0.07	0.01	0.02
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.22: Sibling emigrated

	(1)	(2)	(3)	(4)	(5)
Treatment individual	-0.013	-0.020	-0.0037		
	(0.022)	(0.032)	(0.030)		
Treatment cluster				0.0088	0.016
				(0.019)	(0.020)
Female	-0.032				-0.023
	(0.021)				(0.022)
Control mean	0.15	0.14	0.16	0.14	0.14
N	1153	498	655	2009	1306
R-squared	0.05	0.07	0.08	0.01	0.02
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.23: Sibling got married

	(1)	(2)	(3)	(4)	(5)
Treatment individual	-0.0038	-0.0020	0.0027		
	(0.022)	(0.034)	(0.029)		
Treatment cluster				0.0046	0.0087
				(0.019)	(0.023)
Female	-0.0044				0.0065
	(0.022)				(0.019)
Control mean	0.16	0.17	0.15	0.15	0.15
N	1153	498	655	2009	1306
R-squared	0.04	0.06	0.07	0.01	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq 0.01^{***}$, $\leq 0.05^{**}$, and $\leq 0.1^{*}$. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

Table A.24: Sibling got hazardous work

	(1)	(2)	(3)	(4)	(5)
Treatment individual	0.0021	0.035	-0.022		
	(0.015)	(0.022)	(0.022)		
Treatment cluster				-0.0012	-0.00069
				(0.018)	(0.018)
Female	-0.0078				-0.022*
	(0.015)				(0.013)
Control mean	0.07	0.04	0.09	0.07	0.07
N	1153	498	655	2009	1306
R-squared	0.07	0.12	0.07	0.00	0.01
Controls	Strata	Strata	Strata	Strata	Strata
Gender	Both	Girls	Boys	Both	Both

Notes: All regressions control for strata fixed effects, which are gender and community in columns 1-3, Region in column 4, and Region and gender in column 5. The regressions in columns 1-3 follow equation (1) and the sample is restricted to individuals in treated communities. In addition, the sample consists of only girls in column 2, and of only boys in column 3. The regression in column 4 follows equation 2 and includes all individuals. The regression in column 5 also follows equation 2 but the sample is restricted to control individuals. P-values are $\leq$ 0.01***, $\leq$ 0.05**, and $\leq$ 0.1*. Robust SE in parentheses in columns 1-3 clustered at the community level in columns 4 and 5.

A.4 Pre-analysis plan

Pre-analysis plan for "Effects of speed-schools in Niger"*

Anne Kielland[a] and Andreas Kotsadam [b]

Abstract

We study the effects of speed-schools for older children in Niger. In this plan we describe the hypotheses to be tested and how they will be tested. The description includes how the variables are coded, how we will deal with missing values, and the specification of the estimation equations. We also conduct a power analysis which suggests that we are able to identify relatively small differences. All deviations from the plan will be highlighted in the final paper.

*The research has been funded by the Norwegian Research Council (number 267550).
a Fafo Institute for Labour and Social Research.
b The Ragnar Frisch Centre for Economic Research, Gaustadalleen 21, N-0349 Oslo, Norway.

1. Introduction

The project is studying the effects of a two-year speed-school intervention in Niger that targets rural out-of-school children between 12 and 14 years of age that have never started or dropped out of primary education. The program covers the first six years of the primary school curriculum in two years. It is intended to be a low cost solution that provides a second chance to vulnerable children. We work closely with an NGO, the Strømme Foundation, that has been a pioneer in the speed-school field. Many other organizations have followed and speed schools are relatively common across Africa.

We study the effects of the speed-schools using a randomized controlled trial (RCT). We carried out a two-stage randomization, one at the community level, resulting in treatment and control communities, and the other at the individual level, resulting in "treatment" and "spillover" children in treatment villages, and "pure control" children in control villages. In this plan we describe how we will analyze the data from the experiment.

2. The field experiment

Speed-schools, also known as accelerated learning programs, are intended to let overaged children catch up on their primary education by offering a compressed syllabus. We study such a program in Niger that compresses grades 1-6 into two years. This program is labelled Stratégie de Scolarisation Accélérée 2 (SSA2). Due to an astonishing need in our areas and the limited spaces in the speed-schools, not all children that need and want a speed-school can attend one. Our main intervention consists of children with similar need randomly being offered a place in a speed-school. Additionally, we investigate the effects of a community receiving a speed-school. The experiment is randomized at two different levels:

Randomization across communities: Our experiment takes place in 85 communities in two regions. 30 communities were randomized to treatment and 55 to control. Initially we wanted to have 60 control communities but this was not possible due to too few eligible children in some communities initially listed. It was decided that there would be 15 treated communities in each of two regions covered (Dosso and Tillaberi) and therefore randomization to treatment and control communities were done within regions. In the control communities we listed 1,430 children and interviewed 1,159 children at baseline. In the follow up we will try to reach also children not interviewed at baseline.

Randomization across individuals: In the treatment areas we listed 1,786 children and interviewed 1,489 children at baseline. The schools wanted around 28 individuals per school and it was suggested that we randomized around 35 individuals to treatment as it was deemed likely that treatment compliance would be imperfect. When more than 45 children were listed in a community, we randomly selected 15 to be in the control group. In communities with fewer than 45 individuals listed we assign all non-treated individuals to the control group. Randomization was stratified by gender in the following way: In areas with more than 30 girls we had a quota of 50 percent girls in the treatment group. In areas with fewer girls on the list we randomly assigned treatment to girls in relation to the share of girls on the lists (35*share of girls). The remaining places were randomized to boys. All girls not receiving treatment were allocated to the control group and we filled the rest of the control group places with boys that were not randomly assigned to treatment.

3. Data and coding of main variables

We conducted a listing exercise to map the need in different communities in our two regions. Children were listed based on the following criteria: household within what is locally described as the "academic distance" from village center/prospective Speed-school location (5 km), age (around 12-14), willingness to go to school, parental support and consent to participate in this as a research project. Based on these lists we selected 85 communities and carried out the randomization.

We then collected baseline data in August to November 2018. Baseline interviews were blinded and treatment status was revealed after data collection in the village was finished. Treatment started shortly after baseline data was collected. The intervention was completed in June 2020 and the second round of data collection will take place during December 2020-April 2021.

Data was collected from community leaders, the household heads of the children on the lists, and the children on the lists. We here describe how we will use the survey data and how the variables will be coded.

Before going through the variables we start by describing some general principles in order to avoid repetition:

Standardized indices: When we say that we will use a standardized index we sum the values of the variables, subtract the mean of the control group and divide by the standard deviation of the control group. The control group refers to control children in treated communities. Sometimes we use other types of indices but then that will be explicitly stated.

Indicator variables: When creating a dummy variable from continuous variables, the following procedure is used unless otherwise stated. We first code all non-responses to missing. The binary variable will then be set in order to create 2 groups, as similar in size as possible, while respecting the order of the answers (e.g. from 1 to 4). For example, if 25% answer 1, 25% answer 2, 25% answer 3 and 25% answer 4, then the binary variable will be equal to 1 if the respondent answers 1 or 2 and zero otherwise. On the other hand, if for example 20% answer 1, 10% answer 2, 5% answer 3 and 65% answer 4, then the binary variable will be equal to one of the respondent answers 1, 2 or 3. We will only consider the answers of the respondents in the control group when creating the coding rules for the binary variables.

3.1 Primary and secondary dependent variables

A goal of this project is to provide a comprehensive picture of the impacts of speed-schools on the children, their families, and the communities. We will therefore collect data on a large number of outcomes and we will report the full breadth of the evidence. That being said, power considerations (see below) forces us to limit the number of main hypotheses tested. We group the outcomes into 6 groups and for each group we will have either an index or a main variable to be tested. We have chosen the following main groups:

1) Schooling and learning outcomes.
2) Marriage and fertility.
3) Support for violence.
4) Self-esteem and wellbeing.
5) Hazardous child labor and time use.
6) Gender attitudes.

3.1.1 Schooling and learning outcomes

The overarching program goal of SSA2 is for out-of-school children to transfer back into lower secondary level of the traditional schooling system in Niger. Lower secondary school is divided into academic and vocational tracks (CEG/CET and CFM).

Our main measure for the educational outcomes is *Starting lower-secondary education*. This variable is coded as zero for everyone answering No on the question “Are you enrolled in school or education program this fall?”. It is coded as 1 if they answer yes and then answer CEG/CET or CFM on the follow up question “What type of education program is that?”. We also replace the variable to be equal to one if the respondent answers lower secondary school on the question on the highest grade level achieved.

We will also create a secondary variable, *Qualified for secondary*, equal to 1 if the child answers yes to the question “Did you do the Government test to qualify for lower secondary this summer?” and answer CEG/CET or CFM on the question “What was the result of the Government test?”. Other secondary outcomes will be years of schooling, highest grade level achieved, and results from reading, writing and math tests. We code the following:

Years of schooling will be coded as the continuous answers to the question, “For how many years did you go to school, not counting pre-school?”. Missing values will be replaced by zero for individuals answering No to the screening question “Have you ever been to school?”

Highest grade level. This variable will be coded as ranging between 1-7 where 7 refers to all levels above 6. Together with years of schooling we can see how much they advance from a year of speed-school since we have both of these variables at baseline as well

We have several variables intended to measure learning. First we code a variable, *Illiterate*, to be equal to one if children answer no on the question “Can you read and write French?”. If they answer yes we code them as zero if the manage to write their own name on a piece of paper we provide them and if they can read the sentence “Fatima likes going to school, but she also likes to help her father”. We also create a *Numeracy index* ranging from 0 to 4 where one point is given for being able to solve each of the problems we give them on addition, subtraction, division, and multiplication.

Concerning life skills, we will look at a composite score for knowledge and practices related to nutrition (food with protein), environmental awareness (disposal of plastic bags), malaria prevention (know about bed nets) and Covid-19 (ability to list 4 preventive measures and 2 vulnerable populations, last washed hands).

We have many school related variables that can be used to investigate compliance in our setting.

We will code a dummy variable *Offered SSA2* to equal 1 if the child answers yes on the question: "After we were here two years ago, can you remember receiving the offer to go to the SSA2 in this village?".

We will also describe the answers to the questions, "Did you start the program right away, when the SSA2 first started?" and "Did you complete the program?". In investigating non-compliance we also have questions about why children did not start and why they quit.

There may also be imperfect compliance because the control group may attend the speed-schools. We will code a variable, *Attended the speed school*, which equals one if the child answers yes on the question "In the past two years, have you been attending any school or education program?" and SSA2 on the question "What type of school or program did you attend?".

3.1.2 Marriage and fertility.

Our main measure will be *Appropriate marriage age* which will be equal to the answer on the question "What do you think would be a good age for you to marry?". Non-numerical answers will be coded as missing.

We will also explore the following variables as secondary variables:

Married: Dummy equal to one if answering yes on "Are you married or engaged?"

Parent: Dummy equal to one if answering yes on "Do you have any children?"

Desired fertility: The numerical answer to "How many children would you like to have?" The value will be set to equal to 15 if the numerical value is above 15. We expect quite a few children to answer that it is in the hands of God or similar fatalistic answers and we will code a dummy variable for fatalistic desired fertility as well.

3.1.3 Support for violence

We hypothesize that randomization to SSA2 reduces participants' own use of violence and also the accept for the use of violence more in general. Participants' own inclination to use of violence is assessed by score on the variable “In general, I am able to resolve my problems without using violence”. We label the indicator “*Own violence*”.

Participant accept for the use of violence more in general is proxied by an index based on the following variables:

In certain cases, it is acceptable to use violence to defend health and security of my family
In certain cases, it is acceptable to use violence to defend my private property
In certain cases, it is acceptable to use violence to defend the honor of my family
In certain cases, it is acceptable to use violence to defend political principles
In certain cases, it is acceptable to use violence to defend religious principles

All variables are scored on a scale from 0 (Strongly disagree) to 10 (Strongly agree). We will dummy code the variables based on our general rules and then create the index indicator “*Violence accept*”.

Niger, like the other countries in the region, is increasingly experiencing armed attacks, notably from misguided individuals citing religious motives, often labelled jihadists. It is worth noticing that the way Islam is practiced in Niger, no matter religious affiliation, inspires moral and self-discipline, not the use of violence.

For the purposes of the study, our interest is if being randomized to SSA2 may lower probability of jihadist sympathies, represented by the combination of religious intolerance combined with an acceptance for violence. With regards to acceptance for violence, our proxy is a continuous variable for *Accept religious violence*: “In certain cases, it is acceptable to use violence to defend religious principles”, which is coded on a scale from 0 (Strongly disagree) to 10 (Strongly agree).

With regards to *Religious intolerance*, we use a continuous variable based on: “If my friend abandons the faith, I will avoid him”, also coded on a range from 0 (Strongly disagree) to 10 (Strongly agree).

Since our religious radicalism is a function of (1) religious intolerance and (2) acceptance of violent and/or illegal means (Ozer and Bertelsen 2018) we create a composite dummy variable equal to 1 if the sum of the two variables is > = 10, and each indicator gets at least a score of 5. To make a more precise proxy for jihadist sympathies we recode to zero respondents who on the question “Why do you think some young people these days chose to join an armed group” select response category a; “to gain personal power”, b; “to enrich themselves” or j; “they are bandits”. Similarly, we code to zero those who, to the question “What should government do to prevent young people from joining armed, violent groups” respond a) prohibit radical preaching. We call this variable *Radical* and it will be our main measure.

We will also explore possible variation in the combination on violence acceptance and high scores on other indicators of religious intensity such as religious identification, number of times prayed, number of days of religious studies, religious studies a s main activity over past 30 days, experience of religion as strength in life, and acceptance of female missioning.

We will also explore the answers to the following questions, which will be dummy coded:

I'd like to have friends with other religions than mine.
If you think about yesterday, how many times did you pray?
Religion and faith are sources of strengths to me
I'd like to have friends with other religions than mine
I'd like to have friends from other ethnic groups than mine
If my friend becomes pregnant outside of marriage, I would avoid her
People who go to live non-muslim countries often get detached from their religion and their traditions

In addition we will look into answers to the questions: “Why do you think some young people these days chose to join armed groups” and “What should government do to prevent young people from joining armed, violent groups”.

3.1.4 Self esteem and wellbeing

Our main variable will be *Self-esteem.* We use a validated French translation of the Rosenberg self-esteem scale (Vallieres and Vallerand 1990) which contains the following 10 survey questions:

I am generally satisfied with myself.
Sometimes, I think I'm no good.
I think a have several good qualities
I'm able to do the same things as other people
I don't have much to be proud of
Sometimes, I certainly feel useless
I feel that I'm as valuable as anyone else
I wish I had more self respect
I generally think that I'm a failure.
I have a positive attitude to myself

The answer alternatives are: 0 Strongly agree, 1 Agree, 2 Disagree, 3 Strongly disagree. Items 2, 5, 6, 8, 9 are reverse scored. We therefore give “Strongly Disagree” 0 points, “Disagree” 1 point, “Agree” 2 points, and “Strongly Agree” 3 points on these items. We then sum the scores for all ten items and keep scores on a continuous scale. Higher scores indicate higher self-esteem. A score below 15 (out of the total 30) is usually interpreted as low self esteem and we will also create a dummy for this.

We will also explore effects on happiness:

Happiness: We will reverse code the answers to the question, How happy are you? (1 Very happy 2 Quite happy 3 Not very happy 4 Not at all happy).

As well as the following variables:

In the past two weeks, how often have you felt a low interest for or little pleasure in doing things?

In the past two weeks, how often have you felt sad, depressed or filled with hopelessness?

Since the last time we were interviewing for this study, two years ago, would you say that your life has become better, worse or is about the same as then?

Today, how well would you rate your life conditions on a scale from 0-10, where 0 is very bad, and 10 is perfect?

In general, how would you describe your life conditions compared to those of other young people in the village?

In general, how would you describe your future prospects compared to those of other young people in the village?

How would you rate your own health on a scale from 0-10, where 0 means very bad, and 10 means excellent?

3.1.5 Hazardous child labor and time use

Our main variable will be a dummy variable, *Hazardous child labor*, which is equal to 1 if the child answers that they during the past two years have worked in at least one of the following types of work places:

A household of someone who were not a relative
A mine
A quarry
The streets in a city
On a farm handling chemicals
On a farm where one could sense the smell of chemicals
Carrying loads that one felt were painfully heavy

We will also code a variable, *Working*, equal to one if the child does answer anything but No to the question “Last week, did you work and get paid in cash or kind?”. The other response categories are: Yes for cash, Yes for in kind, or Yes for both cash and in kind payment.

We will also create other types of work variables by investigating answer to the question “What would you say was your main activity in the past 30 days?”. In particular we will create dummies for domestic work, farm work, market work, business, other street work.

Acknowledging that most child labor in Africa is non-remunerated, we will also explore effects on time use by investigating effects on the different time use variables that come after the following in the survey: “We would like to learn about how you spend your time. If you think about the 7 days of the past week: I will read out some activities, and you can tell me if you have done that during the last week, and if so, how many times”.

3.1.6 Gender attitudes

We have a section in the survey with attitude questions after the prompt: “I will read out some different statements to you, and I would like for you to tell me if you agree or not. You should answer exactly what you'd like. There are no right or wrong answers.” The answer categories are from 0 to 10 where 0 is strongly disagree and 10 is strongly agree.

Our main outcome will be a gender equality index based on dummy coding of the following variables:

It is more important to send boys to secondary school
Girls should be allowed to study in higher secondary school even if it is far away
When remunerated work is rare, men should have priority to the jobs

The dummy codings follow our general rule for such codings where we split the variables so that we get as equal sized groups as possible but we will code the dummies so that 1 is always the more gender equal answers.

We will also explore answers to the question about gendered social norms regarding moving away to study. This is based on the question “In this community, most people think girls should be allowed to study at the higher secondary level, even if it is far away”

We will create a variable, *Share of cross sex friends,* with the help of the following two variables: “How many friends do you have, that you would say are close friends?” and “How many of those close friends are girls?”.

In addition we will explore answers to the questions “In certain cases, it is justifiable for a man to beat a woman”, and “Women should be allowed to leave their villages to go on mission/Spread the word of God.”

3.2 Main independent variables from the baseline survey

All main independent variables are from the baseline surveys. The value codings and response categories are often different from baseline to endline but as we use ancova specifications (see below) and have general coding rules for dummy variables, this is not a problem.

We will always have the strata variables (female and community fixed effects for the individual analysis and region fixed effects for the community analysis) as independent variables.

The other main independent variables are:

<u>Continuous numerical values of:</u>

Years of schooling: Already defined.

Desired fertility: Already defined.

Self-esteem: Already defined.

HH respondent age: Numerical value of the age of the household respondent.

Asset index: Based on the principal component analysis of the following items:

Does the household have/own… (recoded as Yes=1, No=0)?

1. *Radio*
2. *Fridge*
3. *Cellphone*
4. *Bicycle*
5. *Moped*
6. *Motorbike*
7. *Car*
8. *Electricity*
9. *Solar panel*
10. *Generator*
11. *Push cart*
12. *Cistern or water reservoir*

The asset index will be standardized by subtracting the mean of the control communities and dividing by the standard deviation of the control communities.

Dummy variables for:

Religious intolerance: "If my friend abandons the faith, I will avoid him/her."

Accept religious violence: "It is acceptable to use violence to defend religious principles."

Gender Equality[1] *school*: "It is more important to send boys to secondary school"

Gender Equality work: "When remunerated work is rare, men should have priority to the jobs"

Gender Equality school: "It is justifiable for a man to beat a woman if she: burns the food"

HH respondent male: Equal to 1 if the household respondent is male.

3.3 Other outcome variables

These variables will be used as exploratory outcomes and to investigate mechanisms.

3.3.1. S*ocial desirability index.*

[1] Note that the gender equality variables are coded such that 1=gender equal.

We will create a *Social desirability index* to investigate whether experimenter demand effects are upward biasing the estimated program impacts, The index is based on responses on a Marlowe-Crowne module which measures a person's general tendency to give socially desirable answers. We use a validated French translation (Daigle 2019). The index is based on the following questions where individuals can answer agree or disagree:

i. It is sometimes hard for me to go on with my work if I am not encouraged.
ii. I sometimes feel resentful when I don't get my way.
iii. On a few occasions, I have given up doing something because I thought too little of my ability.
iv. There have been times when I felt like rebelling against people in positions of authority even though I knew they were right.
v. No matter who I'm talking to, I'm always a good listener.
vi. There have been occasions when I took advantage of someone.
vii. I'm always willing to admit it when I make a mistake.
viii. I sometimes try to get even rather than forgive and forget.
ix. I am always courteous, even to people who are disagreeable.
x. I have never been irked when people expressed ideas very different from my own.
xi. There have been times when I was quite jealous of the good fortune of others.
xii. I am sometimes irritated by people who ask favours of me.
xiii. I have never deliberately said something that hurt someone's feelings.

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

For some of the children we have administrative data on the results from the Government test to qualify for lower secondary school. We will compare these to the answers the children give on *Qualified for secondary* and also relate the differences to the social desirability index.

In addition, there are some variables that are extra interesting to see how they correlate with social acceptability, such as "In general, I am able to resolve my problems without using violence" and the attitudes toward domestic violence.

3.3.2 *Migration*

We will measure *Intend to migrate*, which captures international migration and will be coded to equal 1 if the child answers yes to "Do you think you will actually move away from the village in the future?" and "Do you think you may travel to far away destinations, such as Europe?".

We will also measure national migration by investigating the answers to the questions:

Do you ever think about that you would like to travel to live outside this vilage?
If you should go live outside this village, where would you like to live? (you can chose several)
Are you—alone or with friends or family—seriously considering migrating to another country?

For those that have already moved we will also describe where they have migrated to the extent possible.

3.3.3. Other outcomes

We will also explore the effects on other variables.

3.4 Variables used to investigate spillovers and for LASSO

Some variables will be used as additional controls when using a LASSO procedure, and some to dig into mechanisms.

3.4.1 *Additional variables to investigate spillover effects*

In investigating spillover effects we will additionally investigate if there are treatment effects on the following variables:

We would now like for you to think about what these siblings were doing in the past two years.
Did any of your siblings start school in the fall of 2019 (last year)?
Was that a boy or a girl or both?
Did any of your siblings quit school in the past two years?
Was that a boy or a girl or both?
Did any of your siblings under the age of 18 leave the village to earn money in the past two years?
Was that a boy or a girl or both?
Did any of your siblings under the age of 18 leave the country during the past two years?
Was that a boy or a girl or both?
Did any of your siblings under the age of 18 marry during the past two years?
Was that a boy or a girl or both?
Did any of your siblings under the age of 18 go to work in a mine or a quarry during the past two years?

Was that a boy or a girl or both?

The answers will be dummy coded. The question about gender will be explored to see if the spillovers are gendered.

In addition we will use the household survey questions about what kids have been doing during the last two years as well as the household questions about gender equality.

3.4.2. *Additional variables to be used in the LASSO control set:*

As there is no limitation on how many variables that can meaningfully be included in the LASSO regressions we will include baseline variables and household variables for all variables described in section 3. In addition we will explore whether the variables on building material for the roof, floor, and walls help us in precision beyond the asset index.

4. Empirical strategy and hypotheses

Our empirical strategy is intended to capture individual treatment effects, spillover effects, and effects at the community level.

4.1 Individual level intention to treat effects of treatment

For the analysis of individual level treatment effects we restrict the sample to the treated communities and estimate the following regression:

$$Y_{2i} = \beta Treated_i + \chi X_{1i} + \varepsilon_i \qquad (1)$$

where 2 indicates follow up and 1 indicates baseline, i indexes individuals. The vector X will always include fixed effects for the strata variables gender and community. In addition we will see if we can improve precision in the estimates by including the controls listed in 3.2 and by picking optimal controls from the total list of controls (see 3.4.2) using LASSO (Belloni et al. 2014; Ahrens et al. 2018). If we have missing values on explanatory variables we will code the variables as zero and include dummy variables controlling for missing status so that we do not lose observations. We use robust standard errors in all estimations unless otherwise stated. Unless there is imbalance across treatment and control (see below), the estimation with only the strata variables will be our main specification.

4.2 Spillovers

The specification in (1) will be a biased estimate of the overall effect in a community if there are spillover effects within villages. The question of spillovers is also interesting in itself. To estimate spillover effects we will restrict the sample to control individuals and estimate the following regression:

$$Y_{2i} = \beta Treated\_Community_i + \chi X_{1i} + \varepsilon_i \qquad (2)$$

where the treatment community is randomly assigned at the community level. As treatment at the community level is randomly assigned within regions we will always include *Region* as a strata variable in these regressions. We still estimate this regression using individual level data as power may be increased by including individual level controls. We do, however, cluster the standard errors at the community level in these regressions. As the sample includes only non-treatment individuals, β identifies within-community spillover effects by comparing control individuals in treatment villages to control individuals in pure control villages. Note that we have not randomly assigned the treatment *share* in different communities so we can not investigate how the spillovers change with different intensities of the treatment without further assumptions (Baird et al. 2018).

We will also investigate the questions described in section 3.4.1 to measure spillover effects.

4.3 Community level effects of schooling

We can also use specification (2) but without excluding any individuals to estimate the total effect of having a speed-school in your area. We can further add a dummy variable for being in the control group and interact Treated_community with being in the control group. This would lead to the following estimating regression:

$$Y_{2i} = \beta T\_Community_i + \eta C\_child_i + \phi C\_child_i * T\_Community_i + \chi X_{1i} + \varepsilon_i \qquad (3)$$

where T is short for Treatment and C is short for control.

The coefficient β now identifies the effect of having a speed-school in the area on the individual kids that were randomized to attend the school. This regression would include region fixed effects, female, and the standard errors would be clustered at the community level.

4.4 Balance and attrition

To test for balance we will regress treatment on the main independent variables described in section 3.2, both individually and together, while controlling for the Strata variables (female, community). We will judge whether the randomization worked by conducting an F-test of whether the control variables jointly predict treatment status.

We will also test for balance in the spillover sample by regressing *Treated_Community* on the control variables one by one and together, again controlling for the strata variables (Region fixed effects in this case). In the balance tests using randomization at the community level we will cluster the standard errors at the community level.

We will probably not manage to reach all the respondents initially sampled. We will check whether attrition and missing outcomes are correlated with treatment. If there are statistically significant differences in attrition or non-response between treatment and control (controlling for the strata variables), we will follow the correction proposed by Lee (2009).

4.5. Instrumental variables

To account for imperfect compliance we will also estimate an IV model of the following form:

$$Education_{2i} = \beta Treated_i + \chi X_{1i} + \varepsilon_i \quad (4)$$

$$Y_{2i} = \text{Predicted}\,(Education_{2i}) + \chi X_{1i} + \varepsilon_i \quad (5)$$

Where we predict education with the randomization and use the predicted values for education in the second stage to calculate the local average treatment effect of education on on our outcomes. Education can be measured as *Starting lower-secondary education* or by *Highest grade level*. We can also use the various learning measures in the first stage to estimate the effects of learning on the various outcomes. These analyses will not be treated as our main analyses but it may be useful to scale the ITT estimates to gain more insight into potential mechanisms.

5. Heterogeneity and further exploratory analyses

We will use machine learning techniques to automate the search for heterogenous treatment effects. There are many different types of machine learning algorithms and we have decided to use the "Generic ML" approach by Chernozhukov et al. (2018). If the omnibus test suggests that there is no heterogeneity we will not present further results from the method. As this field is moving rapidly, however, it is possible that there will be other techniques that are relevant for us once we start analyzing the data.

We will also explore heterogeneity directly with respect to some aspects and we will then estimate the following equation:

$$Y_{2i} = \delta Z_{1i} + \beta Treated_i + \lambda Treated_i * Z_{1i} + \chi X_{1i} + \varepsilon_i \qquad (6)$$

Z= Variables in 3.2. For some outcomes, such as the gender attitudes, it will also be interesting to explore heterogeneity across the dimension of household level gender inequality.

6. Power calculation, reporting, and discussion of null findings

Assuming a sample of 1300 individuals in the treatment areas and 80 percent power we can detect effects as small as 0.16 standard deviations.

We will also adjust the p-values for the fact that we are testing the impact on 6 outcomes. We follow the recommendations of Fink, McConnell, and Vollmer, (2014) and use a method developed by Benjamini and Hochberg, (1995) and Benjamini and Yekutieli, (2001) to minimize the false non-discovery rate. The main advantage of the method is that it is limiting the risk of false discoveries while only adjusting the critical values based on other true hypotheses. The false discovery rate method developed by Benjamini and Hochberg (1995) implies that the m p-values of the i hypotheses are ordered from low to high and that the critical value of the p-value is then p(i) = a*i/m. To illustrate, with 6 hypotheses and a significance level (a) of 0.05, the critical p-value would be 0.0083 for the one with the lowest p-value (0.05* 1/6, which is the same as a Bonferroni correction). For the second hypothesis, the critical p-value is 0.017 (0.05*2/6) and for the sixth it is 0.05 (0.05*6/6). As we have 6 main hypotheses, the most significant effect would have to have a p-value of less than 0.0083 and the minimum detectable effect becomes 0.19 with 1300 individuals (0.18 with 1500 individuals).

Note that we apply this correction to the main dependent variables only; when discussing individual variable results within particular outcome groups, we use conventional significance levels. We use this approach because the purpose of studying individual variables within the outcome groups is to understand mechanisms, rather than to single out particular variables for general conclusions.

We may not publish all of the findings in the same paper but we will create a working paper where we present all results outlined in this plan. This working paper will then be referred to and we will mention all main outcomes studied in each published paper.

Some of our findings are likely to be null results. It is often difficult to judge whether such results are showing a meaningful lack of effect or whether they arise due to low power. To investigate if the effects are meaningful null findings we will conduct equivalence tests with two one-sided t-tests (TOST) and show how large positive and negative effects we can reject. The tests are one sided in equivalence testing as one tests whether effects are larger than a highest value and lower than a lowest value. In practice, the procedure is equivalent to presenting the bounds of a 90 percent confidence interval.

7. IRB approval

This study was reviewed and approved by the Norwegian Centre for Research Data (number 386893).

8. Archive and sharing of replication data

The pre-analysis plan is archived before any follow-up data is collected. We archive it at the registry for randomized controlled trials in economics held by The American Economic Association: https://www.socialscienceregistry.org/ on December 10, 2020. We will start the data collection later in December and it will be ongoing until April 2021.

References

Ahrens, A., Hansen, C. B., & Schaffer, M. (2018). PDSLASSO: Stata module for or post-selection and post-regularization OLS or IV estimation and inference.

Baird, Sarah, J. Aislinn Bohren, Craig McIntosh, and Berk Özler. "Optimal design of experiments in the presence of interference." *Review of Economics and Statistics* 100, no. 5 (2018): 844-860.

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

Benjamini, Yoav and Yosef Hochberg (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing.” In: *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300.

Benjamini, Yoav and Daniel Yekutieli (2001). “The control of the false discovery rate in multiple testing under dependency.” In: *Annals of statistics*, pp. 1165–1188

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1-C68.

Daigle, A. (2019). *Impact des représentations du vieillissement et du sentiment de contrôle sur l'ajustement à la retraite dans une population de nouveaux retraités* (Doctoral dissertation, Université du Québec en Outaouais).

Lee, David S. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76, no. 3 (2009): 1071-1102.

Ozer, Simon, and Preben Bertelsen. (2018) "Capturing violent radicalization: Developing and validating scales." *Scandinavian Journal of Psychology*.

Vallieres, E. F., & Vallerand, R. J. (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi de Rosenberg. *International journal of psychology*, *25*(2), 305-316.