# Supporting Learning In and Out of School: Experimental Evidence from India*

Martina Björkman Nyqvist† Andrea Guariso‡

June 30, 2022

**Abstract**

This paper studies an education program targeting primary school children in rural India, which combines a standard in-school pedagogical intervention with an out-of-school study group program managed by parents. We rely on a cross-cutting experimental design across 200 villages and find the full program to significantly increase children's test scores in mathematics and language by 0.09 and 0.11 standard deviations respectively. When the two program components are implemented in isolation, there is no impact on children's learning. The cost-effectiveness analysis highlights high returns from adopting a multidimensional approach that supports children's learning processes both in and out of school.

*JEL classification:* C93, I21, O1

*We are grateful for comments and suggestions by Abhijeet Singh, Andy de Barros, Timothée Demont, Paul Glewwe, Selim Gulesci, Gregory F. Veramendi, and seminar participants at University of Munich (LMU), UC Merced, University of Goettingen, Aix-Marseille School of Economics (AMSE), University of Bologna, University of Milan-Bicocca, Trinity College Dublin, MWIEDC, EOS Development Webinar, YoDEV Webinar, as well as Fadhil Nadhif Muharam, Artur Obminski, Merve Demirel, and Harry Humes for invaluable research assistance. We also thank J-PAL South Asia and its staff, specifically Srijana Chandrashekhar; Mrignaina Tikku; Dwithiya Raghavan; Tanima; Akanksha Saletore; Maaike Bijker; Rithika Nair and Shobini Mukerji for support and management of the data collection and fieldwork. Finally, we thank Rukmini Banerji and Saveri Kulshreshth at Pratham India and Ingrid Eelde Koivisto at Pratham Sweden for insightful discussions about the programs and for their collaboration throughout the study, as well as for sharing administrative data. All mistakes are our own. Financial support from Carl Bennet AB, the Swedish Research Council (2016-05615) and the Mistra Foundation is greatly appreciated. Andrea also acknowledges funding from the TCD Arts and Social Sciences Benefactions (ASSB) Fund. The study was registered with the AEA RCT registry (0002817) and received ethical approval from IFMR Human Subjects Committee (IRB00007107) and TCD Ethic Review Board (05062018).

†Stockholm School of Economics and Misum.

‡University of Milan-Bicocca.

# Introduction

The dramatic gains in school enrollments that have been achieved in low- and middle-income countries over the past two decades have not been mirrored by improvements in learning levels (World Bank, 2018; Altinok et al., 2018; Le Nestour et al., 2021). India is a case in point: more than 96% of all children in the age group of 6-14 years are enrolled in school, but the share of children in grade 5 that can read a grade 2 text or solve a grade 2 subtraction problem has stagnated below 50% for more than a decade (ASER, 2018). The Covid-19 pandemic and associated school closures further slowed learnings (Moscoviz and Evans, 2022; Björkman Nyqvist and Guariso, 2022), making the quest for effective approaches to boost students' achievements in low-income settings an even more pressing priority. Most programs studied so far in the literature focus on how to make schools more effective and the time spent in school more productive (see Muralidharan (2017) for a recent review). The role of further learning *outside* of the classroom has instead received far less attention.

This paper presents experimental evidence of a primary education program aimed at improving children's learning by combining a community-managed out-of-school program with a standard in-school pedagogical program. The out-of-school component is a *Study Groups* program, managed by community volunteers, which brings together primary school children of different ages to study and learn in groups outside of school hours. The in-school pedagogical program consists of *Learning Camps*, which are structured around the "Teaching at the Right Level" approach: students are temporarily re-arranged in the school based on their actual knowledge rather than on their grade, and teaching is tailored to their level.

We study the program's impact on children's learning using a randomized controlled trial with factorial design across 200 villages in Assam, northeast India. The villages were randomly assigned to one of the four study arms of equal size: full program (with both Study Groups and Learning Camps), Study Groups only, Learning Camps only, or control group. We conducted two main rounds of data collection: prior to the start of the program (mid-2018) and roughly 16 months after the program started (end of 2019). We surveyed the primary public schools located in each village (N=200), a representative sample of children enrolled within each school (N=5,726) and a representative sample of their caregivers (N=4,592).

We find that the full education program improved children's learning in math and language on average by 0.09 and 0.11 standard deviations (SD), compared to children in the control group. These effects translate into an increase in the share of children that achieve minimum standards (i.e. grade-2 knowledge) in mathematics and language, by 20% and 13% respectively. We also investigate the impact of direct exposure (attendance) to the program, using an instrumental variable approach (under the assumption that the program had no other effect on learning than through direct participation), and estimate learning gains between 0.38 and 0.48 SD.

We find no impact on children's learning levels when the two program components were implemented in isolation. The lack of impact of the in-school component (Learning Camps) is in contrast with previous assessments of the same pedagogical approach in Uttar Pradesh, where it resulted in large improvements in childrenâs learning (Banerjee et. al., 2016). We combine our data with data from the Uttar Pradesh study and can rule out a number of potential explanations for this difference in impact. However, we cannot directly identify the key driver(s), which highlights the challenges of translating successful programs into new contexts and over time.

Cost-effectiveness analysis shows that the full program, with both the in-school and out-of-school component, costs between 15 and 18.3 US$ per student per 0.1 SD increase in learning, which puts it in the middle of the cost-effectiveness distribution of 27 education programs assessed by J-PAL (Bhula et al., 2013). The analysis highlights that in the context of the ongoing expansion of the Teaching at the Right Level model across India and Africa, adding the out-of-school component can lead to significant gains at a very low additional cost.

Further analysis shows that timing of the in-school pedagogical intervention matters: by exploiting random variation in when the schools hosted the Learning Camps, we find that children exposed to camps earlier display higher test scores by the end of the study, suggesting that the camps put them on a steeper learning trajectory. We also find that all interventions increase parental engagement through more frequent interactions with the school, while the overall amount of time spent with children remains unchanged. At the school level, we find instead evidence of a substitution effect: schools that hosted the learning camps had a drop in educational inputs (both in terms of investments and interactions with stakeholders).

The existing body of evidence has mostly focused on in-school programs and highlights targeted pedagogical interventions as one of the most effective approaches (Glewwe et al., 2009; Banerjee et al., 2016; Glewwe and Muralidharan, 2016; Muralidharan, 2017). In our context, we find that the pedagogical intervention is only effective when paired with the out-of-school community program. Overall, the literature has paid little attention to continued learning outside of schools, despite the potential complementarities with what happens inside the school. There are few studies that look at the impact of after-school remedial education (Lakshminarayana et al., 2013; Banerjee et al., 2016; Chiplunkar et al., 2021; Romero et al. 2021), in some cases with the support of computer-assisted learning (Linden, 2008; Lai et al., 2015; Muralidharan et al., 2019). While results appear generally encouraging, most of these programs are implemented in schools right after (or before) official school hours and are often led by teachers or specifically trained personnel. The program studied in this paper differs from these previous studies, as it combines an in-school and an out-of-school component, and the out-of-school component is implemented in the community and managed by parents.[1]

Another strand in the literature has studied the impact of parental engagement with preschool children and find large impacts on child development (Attanasio et al. 2020a; 2020b). There is however limited evidence of parental and community engagement with primary school children outside of school hours and the impact it might have on their learning levels (Beasley and Huillery, 2017; Banerji et al., 2017; Islam, 2019; Di Maro et al, 2020). Our study contributes to this literature by showing that parental engagement in children's learning can play an important role in improving educational program effectiveness in a cost-effective way.

# 1. Intervention, sample, evaluation design

The setting for this study is Nagaon district in the state of Assam in India (Figure A.1 in Appendix). According to ASER (2018), school enrolment for children 6-14 years in Assam is nearly universal (97.7%). Yet, learning levels remain low: only 34% of children in grade 5 can read a grade 2-level text and only 14% can solve simple division problems, which ranks Assam well below the national average, where the corresponding shares are 44% and 23% respectively. Nagaon district performs on par with the state average in terms of schooling outcomes (ASER, 2016).[2]

[1] In a recent paper de Barros et al. (2021) study the impact of a teachers' training program, both alone and in combination with an after-school community-led student contests. They find the community-led component to have a negative effect on instruction quality, and no impact on studentsâ learning.

[2] Table A.1 in Appendix compares our study sample across both Nagaon district and the broader state of Assam.

## 1.1. Intervention

The program studied in this paper target primary school children and includes two components, the Study Groups and the Learning Camps, implemented by Pratham Education Foundation.

**Study Groups** The community-based *Study Groups* provide a framework for the community to support out-of-school studying and learning, while at the same time maintaining costs and time commitment low, by harnessing the power of peer support. Pratham introduced the Study Groups in the target villages at the beginning of the school term and mobilized community volunteers, who then locally managed it. Each study group of about six primary school-age children was coordinated by one volunteer (typically a mother) and the total number of groups in the village varied depending on the number of children and volunteers. The structure of the study groups were flexible and it was up to each volunteer to set the frequency and length of the meetings. Pratham only conducted short monthly visits to the villages to distribute self-explanatory learning material that could be used to guide activities in the groups, while the groups were also encouraged to focus on other reading and studying activities, such as completion of school homework.[3]

**Learning Camps** The *Learning Camps* were implemented by Pratham in the schools during regular school hours and consisted of intensive bursts of teaching and learning activities focused on the foundational skills for reading and arithmetic. The camps were based on Pratham's pioneered pedagogical approach called Teaching at

[3]Study groups did not maintain any formal records of their activities but at endline we conducted a short qualitative survey with 40 volunteers across the study villages. There was significant heterogeneity, but most groups met 2 or 3 times per week for about 2 hours, and children mostly focused on the material shared by Pratham.

the Right Level (TaRL), which provides teaching that is tailored to a childâs individual learning level. Pratham introduced the learning camps in the target villages at the beginning of the school term and mobilized community volunteers to help with the implementation. The camps were then administered by Pratham staff with the support of available teachers and volunteers. The camps consisted of three periods of 10 days each (for a total of 30 days), spread over one teaching term (5 months). At the beginning of the camp, Pratham tested all children in grades 1 through 5, to identify their level of reading and arithmetic. Children were thereafter grouped according to their level and learned language and mathematics through specific activities and materials tailored to each group.

**Full program** The full program combines both the Study Groups and the Learning Camps components.

## 1.2. Sample and experimental design

The study sample of 200 villages was randomly selected from a list of villages in Nagaon district that Pratham identified as eligible.[4] Each village hosts one primary public school.

After baseline data collection, the 200 villages were randomly divided into four groups of equal size: *Learning Camps & Study Groups*; *Learning Camps*; *Study Groups*; *Control*. A second randomization was conducted to determine the timing of the implementation of the Learning Camps, given that Pratham could not run them at the same time across all villages due to resource constraints. Out of the 100 villages assigned to receive the Learning Camps (either alone or in combination with the

[4]Pratham's standard preliminary assessment considers official enrollment in the local primary school, accessibility of the village, and a qualitative assessment of the potential for mobilizing the community. Figure A.1 shows the spatial distribution of the study villages within Nagaon district.

Study Groups), we randomly selected (stratifying by group) half of them to receive the program right after baseline data collection (*phase I*), and the other half to receive it roughly 5 months later (*phase II*). Since the endline survey took place at the same time across the whole sample, this introduced random variation in children's time since exposure to the Learning Camps. Figure 1 illustrates the experimental design.

### 1.3. Data

Data for the analysis is collected from three sources: school, children, and their primary caregivers. The school survey included a combination of direct questions to the head teacher and observational records of school infrastructure. The child survey was administered to a random sample of 8 students in each grade 1 to 4 at baseline (thus expected to be in grades 2-5 at endline), for a total of 32 students per school.[5] Student selection was based on official school enrollment registries and sampled students who were absent from school on the survey day were surveyed at home. The child survey consisted of two parts: a test and a short survey. The test component included two tests, each one with a mathematics and a language section. The first test (ASER test) mirrored the standard test conducted yearly by the ASER Center across India for children aged 5 to 16. The second test (Test A) was instead created by the research team based on questions used in other studies in India (Muralidharan et al. 2020) and was designed to provide an alternative to the standard ASER assessment. Both ASER and Test A were conducted orally and individually with each sampled child by trained enumerators (see Appendix B for details). By endline Test A proved to be too easy for the students in the sample and displayed little variation. Therefore, in our main analysis, we focus on the standard ASER test

[5] Because some schools were smaller than the target, we eventually surveyed on average about 7 children per grade or 29 children per school.

and report the results on Test A (as well as on a comprehensive measure including both tests) in the appendix.[6] The student survey mainly contained questions about the childâs studying and learning activities. Finally, at baseline we randomly selected 6 children per grade (i.e. 24 per school) from our sample and tracked their primary caregiver (i.e. the person who dedicates the most time to looking after the needs of the child) at home.[7] The survey collected basic information on household characteristics and education practices.

The baseline survey was conducted between May and August 2018. Implementation of the programs started immediately after baseline data collection. The implementation of Learning Camps was rolled out in two phases, while the community Study Groups activities started in all villages immediately after baseline and continued until the end of the study. A short compliance survey was conducted in May and June 2019. The endline survey was conducted between November 2019 and January 2020, 16 months after the start of the program.

## 1.4. Baseline balance and validation

The schools in our sample had on average 53 students enrolled, 2 classrooms, and 4 teaching staff at baseline. Attendance on survey day was on average 70%. About one in four students (25%). reported never studying after school. In terms of learn-

[6] ASER test has the advantage of being comparable to national tests as well as to previous studies in the literature. One concern is that Pratham's Learning Camps, by dividing students according to the results on an initial ASER-style test, might be especially good in teaching to the ASER test. However, the lack of impact in the study arm exposed to Learning Camps only alleviates this concern. As reported in Appendix, our results are confirmed when considering the Test A mathematics scores, while they are typically non-significant when considering the Test A language scores, where half of the students correctly answering 90% of the questions (the average was 76%). Finally, results are confirmed when we combine both our tests in a single measure, using a two parameter logistic (2PL) item response theory (IRT) model (Jacob and Rothstein, 2016).

[7] The respondent was typically the mother (85%), followed by the father (5%), and the grandmother (3%). Because some schools were smaller than the target, we surveyed on average 23 primary caregivers in each village (see Table A.2 in appendix)

ing levels at baseline, the average student in our sample scored between level 1 (beginner) and level 2 (corresponding to 1-digit number recognition in math and letter recognition in language) in the ASER test components. While there is a progression in learning across grades (Figure A.3 and A.4 in Appendix), the overall learning levels are low: about 35% of students in grade 4 were unable to solve subtractions and 66% were unable to read a story, which are the expected standards for a grade 2 student. Table A.1 in Appendix A shows that observable characteristics of schools, children, and primary caregivers were balanced across study arms at baseline.

At endline we were able to track back 93% of the baseline sample. Attrition rates did not differ across study arms and there was no difference across treatment arms in characteristics of children lost at endline (Table A.3 in Appendix A).[8]

Program implementation and study design compliance are checked in two ways. First, we matched our sample of schools and children with detailed Pratham's administrative records to confirm that Study Groups and Learning Camps took place in every study village. We recorded correct adherence to the original design, with the exception of two villages with similar names, originally assigned to the control and the full program intervention, which had been mistakenly swapped by Pratham. In the analysis, we adopt a conservative approach and rely on the original random assignment.[9] The records show that Study Groups villages had on average 5 study groups and that within Learning Camps villages 93% of the children in our sample attended at least one day of Learning Camp in the schools (with an average attendance of 20 days). Second, at the endline we asked head teachers, children, and caregivers about interactions with Pratham and exposure to the programs. Such

[8]At endline one head teacher from the Study Groups treatment arm refused to answer the endline questionnaire (while still giving permission to test and survey children in the school) and we therefore miss one school survey. Caregivers' attrition mirrors children's attrition (93%).

[9]Using the actual assignment leaves in any case our results virtually unchanged.

recall information is likely noisy, but we find the distribution of the replies across study arms to mirror the study design (Table A.5 and A.6 in Appendix).

Overall, our checks on baseline balance, attrition, and implementation alleviate possible concerns related to the integrity of the study design and hence, differences in outcomes at endline should be attributed to the program.

## 1.5. Empirical Strategy

Our main estimating equation is:

$$Y_{i,v} = \alpha_1 LC\&SG_v + \alpha_2 SG_v + \alpha_3 LC_v + \Omega X_{i,v} + \epsilon_{i,v} \tag{1}$$

where $Y_{i,v}$ is the outcome of child $i$, attending school in village $v$. $LC\&SG$, $SG$, and $LC$, are indicator variables that take value one if the village was assigned to the full program (i.e. both Study Groups and Learning Camps), to the Study Groups alone, or to the Learning Camps alone, respectively. $X_{i,v}$ is a vector of individual-level control variables: age, gender, indicators for grade, and baseline value of the outcome variable $Y_{i,v}$. Standard errors are clustered at the village level. School and household level analysis follows a similar specification, with variables defined at the corresponding level. Finally, to form a judgment about the impact of the intervention on a family of $n$ related outcomes, and address potential multiple hypothesis testing concerns, we follow Kling et al. (2004) and estimate a seemingly unrelated regression system to derive the average standardized treatment effect (ASTE).

# 2. Results

## 2.1. Student Learning Outcomes

Table 1 presents our main results of the impact of the program on children's learning outcomes. We find that the full program significantly increased mathematics and language test scores by on average 0.11 and 0.09 SD, respectively, compared to the control group.[10] The average effect size of 0.1 SD corresponds to the median effect size found across 270 different educational programs implemented in low- and middle-income countries, recently reviewed by Evans and Yuan (2021). In our context, it translates to an increase of 20% (13%) of primary school children that achieve minimum standards (i.e. grade 2 level) in mathematics (language) (columns 2 and 4).

When studying the impact of the two program components (Learning Camps in the schools and Study Groups in the villages) implemented individually, we do not find any increase in children's test scores, irrespective of the subject. Not only are the results not statistically significant, but the estimated coefficients are close to zero. The p-values confirm that the full program led to a significantly higher impact on children's test scores than either of the two components alone, with the differences being significant at the 10% level.[11]

The lack of impact of the Learning Camps component alone is in contrast with the results in Banerjee et al. (2016), which look at the same pedagogical model in Uttar Pradesh. To investigate this further, we combine their data with ours.[12] The

[10]We follow the literature (e.g., Banerjee et al., 2016) and normalized ASER test scores using the mean and standard deviation for the control group.

[11]Figure A.5 in Appendix A shows the change in learning levels experienced across the different study arms, while Figure A.6 and A.7 report the endline distribution of the different learning levels by grade and study arm.

[12]We thank the authors for making their dataset available to us.

large difference in the impact of the Learning Camps across the two settings (.58 SD in math and .69 SD in reading) remains unchanged even after controlling for a rich set of potential explanatory variables (test score starting level, age, grade, gender, whether the student was surveyed in school or at home, the time between surveys, school size, as well as their interactions).[13] This suggests that the different impact we observe depends on dimensions not captured in the data. And while our study shares many similarities with Banerjee et al. (2016) (pedagogical model, implementing partner, data collection agency, assessment tool), there are still many differences in terms of setting and program details (see Table A.8 in Appendix for a comparison of the two studies). Overall, this highlights the importance of continuous program monitoring and adaption to the local contexts as a central feature in any scale-up activity.

As already mentioned above, not every student in the treatment arms participated in the program activities. In columns 5 to 8 in Table 2 we thus move to estimate the impact of *direct* participation, by relying on an instrumental variable approach. We compute exposure to the Learning Camps by using Pratham's detailed attendance sheets[14], while we estimate direct exposure to the Study Groups using children's self-reported information on having attended a study group in the village.[15] We then estimate the 2SLS estimates using the following equation:

[13]Results are reported in Table A.9 in Appendix. The sample includes only students enrolled in grades 3 and 4 at baseline, to ensure comparability across the two studies.

[14]Out of the 2,760 children in our Learning Camps sample, 93% attended at least one day of Learning Camps. The share of children in our dataset that we successfully matched with Pratham's administrative data was similar across LCSG and LC groups (p-value=0.717). Average attendance was 64% of the days and the distribution was similar across Learning Camps only and full program schools, as shown in Figure A.10 in Appendix.

[15]The study groups were managed by the community without any close monitoring or supervision and without any administrative record. On average, the share of students reporting that they attended study groups at least once is 17% in the Study Groups only study arm, 24% in the full program study arm, and 4% in the other two study arms.

$$(2) \qquad Y_{i,v} = \beta_1 \widehat{LC_{att}\&SG}_{att\,i,v} + \beta_2 \widehat{SG_{att}}_{i,v} + \beta_3 \widehat{LC_{att}}_{i,v} + \Omega X_{i,v} + \varepsilon_{i,v}$$

Where $\widehat{LC_{att}\&SG}_{att}$ takes value one if student $i$ attended at least one day of the Learning Camp in the school and reported attending at least once the Study Groups in the village, $\widehat{SG_{att}}$ takes value one if the student reported attending the Study Groups, but never attended the Learning Camps, and $\widehat{LG_{att}}$ takes value one if the student attended the Learning Camps, but reported never attending the Study Groups. We instrument the three endogenous attendance variables with the random assignment to the different study arms.[16] The 2SLS estimates in Table 2 (columns 3 to 4) indicate that the program had a large impact on children who attended both the Learning Camps in the schools and the Study Groups in the villages, with learning gains ranging between 0.38 and 0.48 SD.

## 2.2. Cost Effectiveness

We use data from Pratham to calculate cost-effectiveness for each of the intervention arms, estimating the cost of serving 50 villages with the full program, Learning Camps only, and Study Groups only. There are a few indications emerging from the cost analysis (Table 2). First, as expected, the Learning Camps component is significantly more expensive (60% more) than the Study Groups component. Second, there are strong cost synergies that can be exploited by combining the two components. The average cost for the Learning Camps only program was 15.9 US$ per student, for the Study Groups only program was 10 US$ per student, and it was 17.2 US$ per

[16]We make the (strong) assumption that the program only affected learning through direct participation. The most likely direction of the potential bias from violating the exclusion restriction is towards underestimating the true effect of direct exposure, as children not directly exposed might benefit from interactions with exposed children.

student for the combined program. This means that in the context of an expansion of the Learning Camps program, which we are currently witnessing across India and the African continent, it costs very little to add the study group components, with potentially very large returns.

Using the learning gains reported in Table 2 and the two cost-effectiveness numbers, we find the cost of increasing learning outcomes by 0.1 SD to range between 15 US$ and 18.3 US$ per student for the full educational program which implies an impact between 0.55 SD and 0.67 SD per 100 US$ investment (under the simplifying assumption of linear returns). To put these numbers in perspective, the program falls right in the middle of the distribution of a set of 27 education programs aimed at improving learning outcomes and evaluated using randomized controlled trials for which J-PAL has assessed the cost-effectiveness (Figure A.8 in Appendix).

## 2.3. Heterogeneity

We randomly assigned schools to host the Learning Camps in different phases, with a 5-month difference in time. Columns 1 and 2 of Table 3 report the results when we estimate equation (1) and interact the indicators for assignment to a Learning Camps study arm (*LC&SG* or *LC*) with an indicator for being in the first implementation phase. The estimates show that children in schools that hosted the Learning Camps earlier achieved higher test scores by endline. Although imprecisely estimated, the results suggest that the full education program led to learning improvements twice as large for students who received the program earlier on.[17] This result is consistent with the dynamic complementarity hypothesis of skill formation (Cunha and Heckman, 2007): Learning Camps managed to put children on a steeper learning trajectory, whereby they could cumulate more knowledge over time and, hence,

[17]Figure A.9 in Appendix visually illustrates this difference.

widen the gap with the control group. The results appear particularly clear for the full intervention, suggesting that the out-of-school intervention provided children with an opportunity to take advantage of the steeper learning trajectory to further strengthen their learnings.

We find instead no differential impact of the program across children's basic characteristics (gender, grade, and starting test score level) (columns 3-8 of Table 3).

## 2.4. Impact on children, schools, and caregivers

In Table 4 we look at a set of additional child, caregiver, and school outcomes. Panel A focuses on child-related outcomes. The programs had no impact on school attendance (column 1), indicating that the impact on test scores is not driven by changes in exposure to regular teaching. Children in the Study Groups villages report a higher likelihood of studying after school (column 2) and of participating in the study groups in particular (column 3), with the share jumping from zero in the control villages to 17% in the Study Groups villages and 24% in the full program villages. We do not find any discernible effect on their aspiration of schooling achievements (which is fairly high, with the median child in our sample reporting that she would like to complete secondary school and 42% reporting that they would like to continue with university education, column 4), or on the children's overall perception of schooling and learning (column 5).[18]

In Panel B we show that caregivers in the Study Groups villages were more likely to be involved in the organization of the study groups (column 1), but none of the interventions affected the amount of time caregivers spent with children (column 2),

[18]To capture aspirations, we asked until which grade the child would like to study. To measure school and learning perceptions we asked three different sets of questions and combined all answers using principal component analysis: 1) the importance they attach to education; 2) how much they like school, reading, and mathematics; 3) a set of questions about perception of schooling and learning.

or the time they spent helping with homework (column 3). There was also no impact on household spending on education (column 4). These results are consistent with the design of the programs, which were meant to keep a low demand in terms of time and money from parents. All three programs, however, significantly increased caregivers' interactions with the school (column 5).[19] Column 6 reports the average standardized treatment effect (ASTE) across all five caregiver variables and show an overall increase in parental engagement in the villages that hosted the programs. This finding is encouraging, as it is notoriously difficult to increase parental engagement in education, even with high-intensity programs (e.g. Di Maro et al, 2020). However, this also indicates that higher parental engagement does not automatically translate into higher child learning, which is also consistent with what for instance Barrera-Osorio et al. (2020) recently found in a more heavy-handed intervention on parental engagement in Mexico.

Finally in Panel C, reporting school outcomes, the results show that the programs had no impact on teachers' attendance (column 1), while we find an overall *negative* impact on both school-level physical inputs (column 2), which include teaching and learning material and investments in the school infrastructures, and on school's engagement with stakeholders, as captured by the frequency of the meetings with school management committees (column 3) and government officials (column 5). We find no negative impact on the frequency of meetings between parents and teachers (column 4). Column 6 reports an overall reduction in school-level inputs especially large for schools that hosted the Learning Camps program, suggesting that the program might have acted as a substitute for the schools' and governments' inputs in the education process. To the extent that these inputs support the chil-

[19]Interaction is measured through an index that combines five variables using principal component analysis: attendance to parent-teacher meetings, discussing education with teacher, receiving information from the school, discussing education with the school, and checking children's marks.

dren's learning process in school, such a substitution effect might have undermined the effectiveness of the program.

## 3. Conclusion

This paper provides the first experimental evidence on the impact of an educational program that combines a community-managed out-of-school component with a standard in-school pedagogical intervention targeting primary school children.

We find that the program was successful in raising children's learning levels in both mathematics and language by 0.1 SD on average, over a period of 16 months. The result shows that the combination of the two components was important: when the programs were implemented alone, neither of them had any impact on children's learning. These findings speak to the multi-dimensionality of the learning process and show that educational programs that intervene on several dimensions at the same time can take advantage of complementarities (also in terms of costs) and lead to significant learning gains for the children. These findings also highlight the important role that out-of-school learning programs can play, despite having received little attention so far both in policy and in the literature. We estimate the full program to be quite cost-effective and the implementing organization, Pratham, recently made it its new flagship educational model across India. Our results also indicate that organizations should test ways to increase program participation, in order to ensure that the benefits of the programs are fully grasped by the target beneficiaries.

We also examine how program impact evolves over time, which is something quite rare in the education literature, where only 1 in 10 studies is assessed more than one month after the intervention ended (McEwan, 2015). We do so both by exploiting random variation in the timing of the implementation within our original sample, finding estimates that are significantly different from our average impact. This shows that short-term evaluations of education programs, which represent the norm in the literature, are likely to provide a very incomplete assessment of the programs on children's outcomes.

The lack of impact of the pedagogical, Leaning Camps program, when implemented in isolation, is in contrast with previous findings in the literature which further highlights the importance of reviewing and adapting even the most successful programs to the local contexts. A recent study by Duflo et al (2021) provides a step in this direction, by testing alternative models of the Teaching at the Right Level approach in Ghana, where they find impacts between 0.08 SD and 0.15 SD. Our results suggest an alternative way to enhance the effectiveness of the traditional Learning Camps model, by combining it with an out-of-school, community-led, study groups program.

Future research should investigate how the NGO and evaluation teams best can work with a program adaptation approach when scaling up programs (such as the Teaching at the Right Level program) in order to tweak and modify the program along the implementation phase to fit the new context. More in general, more research is needed to study out-of-school programs where parents and community are engaged in the children's learning process. Finding effective ways to improve uptake and engagement of the community and parents to increase local ownership of these initiatives and facilitate their scaling up is an important next step.

# References


[1] **Altinok, N., Angrist, N., and Patrinos, H. A.** 2018. Global data set on education quality (1965-2015). *The World Bank Policy Research Working Paper* 8592.

[2] **ASER.** 2016. *Annual Status of Education Report (Rural) 2016*. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf, Pratham Organization.

[3] **ASER.** 2018. *Annual Status of Education Report (Rural) 2018*. http://img.asercentre.org/docs/ASER%202018/Release%20Material/aserreport2018.pdf, Pratham Organization.

[4] **Attanasio, O., Meghir, C., and Nix, E.** 2020a. Human capital development and parental investment in India. *The Review of Economic Studies*, 87(6): 2511-2541.

[5] **Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M.** 2020b. Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1): 48-85.

[6] **Banerjee, A., Banerji, R., Duflo, E., Glennerster, R., and Khemani, S.** 2010. Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1): 1-30.

[7] **Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M. and Walton, M.** 2016. Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India. *NBER working paper series* w22746.

[8] **Banerji, R., Berry, J., and Shotland, M.** 2017. The impact of maternal literacy and participation programs: Evidence from a randomized evaluation in India. *American Economic Journal: Applied Economics*, 9(4): 303-37.

[9] **Barrera-Osorio, F., Gertler, P., Nakajima, N. and Patrinos, H.** 2020. Promoting Parental Involvement in Schools: Evidence From Two Randomized Experiments. *NBER working paper series* w28040.

[10] **Beasley, T. E. and E. Huillery.** 2017. Willing but Unable? Short-term Experimental Evidence on Parent Empowerment and School Quality. *World Bank Economic Review*, 31(2): 531-552.

[11] **Bhula, R., Mahoney, M. and K. Murphy.** 2013. Conducting cost-effectiveness analysis (CEA). *J-PAL publication*, available at https://www.povertyactionlab.org/resource/conducting-cost-effectiveness-analysis-cea.

[12] **Björkman Nyqvist, M. and A. Guariso.** 2022. The impact of the COVID-19 pandemic on children's leaning and wellbeing: Evidence from India. *Mimeo*.

[13] **Chiplunkar G., Dhar D., and R. Nagesh.** 2021. Too little, too late: Improving post-primary learning outcomes in India. *Mimeo*.

[14] **Cunha, F. and J. Heckman.** 2007. The Technology of Skill Formation. *American Economic Review*, 97(2): 31-47.

[15] **de Barros, A., Fajardo-Gonzalez, J., Glewwe, P., and A. Sankar.** 2021. Public Servants Under Pressure: Experimental Evidence on Efforts to Improve Teaching in India. *Mimeo*.

[16] **Di Maro, V., Leeffers, L., Serra, D. and P. C. Vicente.** 2020. Mobilizing Parents at Home and at School: An Experiment on Primary Education in Angola. **Novafrica Working Paper** No 2002.

[17] **Duflo, A., J. Kiessel and A. Lucas.** 2021. Experimental Evidence on Alternative Policies to Increase Learning at Scale. *NBER Working Paper* w27298.

[18] **Evans, D., and F. Yuan.** 2021. What We Learn about Girls' Education from Interventions that Do Not Focus on Girls. *Mimeo*.

[19] **Glewwe, P., Kremer, M., and Moulin, S.** 2009. Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, 1(1): 112-35.

[20] **Glewwe, P., and K. Muralidharan.** 2016. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, Vol. 5, pp. 653-743. Elsevier.

[21] **Islam, A.** 2019. Parent-teacher meetings and student outcomes: Evidence from a developing country. *European Economic Review*, 111: 273-304.

[22] **Jacob, B., and J. Rothstein.** 2016. The Measurement of Student Ability in Modern Assessment Systems. *Journal of Economic Perspectives*, 30 (3): 85-108.

[23] **Kling, J.R., Liebman, J. B., Katz, L. F., and L. Sanbonmatsu.** 2004. Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment. *Princeton University Industrial Relations Section Working Paper* 481.

[24] **Lai, F., Luo, R., Zhang, L., Huang, X., and S. Rozelle.** 2015. Does computer-assisted learning improve learning outcomes? Evidence from a randomized experiment in migrant schools in Beijing. *Economics of Education Review*, 47: 34-48.

[25] **Lakshminarayana, R., Eble, A., Bhakta, P., Frost, C., Boone, P., Elbourne, D., and V. Mann.** 2013. The Support to Rural India's Public Education System (STRIPES) trial: A cluster randomised controlled trial of supplementary teaching, learning material and material support. *PloS one*, 8(7), e65775.

[26] **Le Nestour, A., Moscoviz, L., and J. Sandefur** 2021. The long-term decline of school quality in the developing world. *Mimeo*.

[27] **Linden, L. L.** 2008. Complement or substitute?: The effect of technology on student achievement in India. *Working Paper, Columbia University: InfoDev*.

[28] **McEwan, P. J.** 2015. Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, 85(3): 353-394.

[29] **Moscoviz, L. and D. Evans** 2022. Learning Loss and Student Dropouts during the COVID-19 Pandemic: A Review of the Evidence Two Years after Schools Shut Down. *Center for Global Development Working Paper* 609.

[30] **Muralidharan, K.** 2017. Field experiments in education in developing countries. In *Handbook of economic field experiments*, Vol. 2, pp. 323-385. North-Holland.

[31] **Muralidharan, K., Singh, A., and A. J. Ganimian.** 2019. Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, 109(4): 1426-60.

[32] **Muralidharan, K. and A. Singh.** 2020. Improving Public Sector Management at Scale? Experimental Evidence on School Governance India. *NBER Working Paper* w28129.

[33] **Romero, M., Chen, L., and N. Magari.** 2021. Cross-Age Tutoring: Experimental Evidence from Kenya. *Economic Development and Cultural Change*, Forthcoming.

[34] **World Bank.** 2018. *School enrollment, primary (% net) - Low & middle income* [Data file.] `https://data.worldbank.org/indicator/SE.PRM.NENR?locations=XO`.

# FIGURES & TABLES

**Figure 1:** Experimental Design

| | | | Study Groups | |
|---|---|---|---|---|
| | | | *No* | *Yes* |
| **Learning Camps** | *No* | | 50 villages | 50 villages |
| | *Yes* | *Phase I* | 25 Villages | 25 Villages |
| | | *Phase II* | 25 Villages | 25 Villages |

Notes: The 200 villages were randomly assigned to one of the four main study arms as shown in the above table. Villages assigned to receive the Learning Camps were further randomized in two groups to determine whether they were to host the Learning Camps in the first or second phase of the implementation.

**Table 1:** Program impact on test scores

| *Model:* | OLS | | | | 2SLS | | | |
|---|---|---|---|---|---|---|---|---|
| *Dep Var:* | Math | | Language | | Math | | Language | |
| | ASER | Grade II | ASER | Grade II | ASER | Grade II | ASER | Grade II |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| LC & SG | 0.11** | 0.06** | 0.09** | 0.04* | 0.48** | 0.23* | 0.38* | 0.20* |
| | (0.05) | (0.03) | (0.04) | (0.02) | (0.24) | (0.13) | (0.20) | (0.12) |
| SG | -0.01 | -0.00 | 0.01 | 0.02 | -0.06 | -0.01 | 0.04 | 0.14 |
| | (0.05) | (0.03) | (0.04) | (0.02) | (0.38) | (0.21) | (0.32) | (0.17) |
| LC | 0.02 | 0.02 | 0.02 | 0.00 | 0.00 | 0.01 | 0.00 | -0.00 |
| | (0.05) | (0.03) | (0.04) | (0.02) | (0.06) | (0.04) | (0.06) | (0.03) |
| Basic controls | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ |
| Mean Control group | -0.00 | 0.30 | -0.00 | 0.31 | -0.00 | 0.30 | -0.00 | 0.31 |
| R-squared | 0.415 | 0.214 | 0.580 | 0.314 | 0.407 | 0.210 | 0.576 | 0.314 |
| Observations | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 |
| No. of clusters | 200 | 200 | 200 | 200 | 200 | 200 | 200 | 200 |
| p-val(LC & SG=SG) | 0.03 | 0.06 | 0.07 | 0.40 | | | | |
| p-val(LC & SG=LC) | 0.08 | 0.15 | 0.10 | 0.13 | | | | |
| F-stat. LC & SG | | | | | 35.53 | 35.49 | 35.83 | 36.18 |
| F-stat. LG | | | | | 22.07 | 22.64 | 22.18 | 22.63 |
| F-stat. LC | | | | | 1,598.51 | 1,574.24 | 1,602.67 | 1,626.61 |

Notes: The dependent variables are children's ASER test score, normalized using the mean and standard deviation for the control group (odd columns), and an indicator for the student reaching at least Grade II level in the ASER test (even columns). In columns 1 to 4 LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both Learning Camps and Study Groups), to the Study Groups alone, or to the Learning Camps alone, respectively. In columns 5 to 8 the variables capture instead direct participation in the program, measured as follows: *LC & SG* takes on value 1 if child attended at least one day of the LC AND reported ever attending the SG; *SG* takes on value 1 if child reported ever attending the SG BUT did not attend any day of the LC; *LC* takes on value 1 if child attended at least one day of the LC BUT reported never attending the SG. These variables are instrumented using assignment indicators to the different treatment arms. *Basic controls* include age, gender, baseline value of the dependent variable and grade fixed effects. The p-values in the bottom of the table are the test of the null hypothesis of equal treatment effects between different intervention arms. F-stats from the first stage are reported at the bottom of the table. Standard errors clustered by village in parentheses. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

**Table 2:** Cost Analysis

| | LC & SG | LC | SG |
|---|---|---|---|
| | (1) | (2) | (3) |
| **Total costs (yearly)** | **34,253** | **31,756** | **19,853** |
| Personnel | 23,685 | 23,685 | 13,293 |
| TLM | 5,369 | 3,454 | 3,398 |
| Training | 2,619 | 2,037 | 582 |
| Travels | 614 | 614 | 614 |
| Other Costs | 1,966 | 1,966 | 1,966 |
| # years | 1.33 | 1.33 | 1.33 |
| # villages served | 50 | 50 | 50 |
| # children served per village | 53 | 53 | 53 |
| Avg cost per student | 17.2 | 15.9 | 10.0 |
| Avg learning gains (SD) | 0.09 - 0.11 | 0.018 - 0.023 | -0.01 - 0.01 |
| **Cost per 0.1 SD gain** | 15.0 - 18.3 | | |
| **Additional SD per 100 US$** | 0.55 - 0.67 | | |

Notes: all measures expressed in US$. *TLM* indicates the Teaching and Learning Material developed and printed by Pratham. *Travels* include the travel cost for Pratham's supervisors, while the cost for individual staff members to reach the different villages is included in their salary. *Other costs* include fixed costs such as office space, utilities, and workshops.

**Table 3:** Heterogeneous Effects

| | ...Phase I | | ...Girl | | ...Grade | | ...BL Score | |
|---|---|---|---|---|---|---|---|---|
| | Math | Lang. | Math | Lang. | Math | Lang. | Math | Lang. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| LC & SG × ... | 0.08 | 0.10* | -0.00 | -0.09 | 0.03 | 0.01 | 0.00 | -0.01 |
| | (0.07) | (0.05) | (0.07) | (0.06) | (0.03) | (0.03) | (0.03) | (0.03) |
| SG × ... | | | 0.02 | -0.07 | 0.01 | 0.02 | 0.01 | 0.04 |
| | | | (0.06) | (0.06) | (0.03) | (0.03) | (0.03) | (0.04) |
| LC × ... | 0.06 | 0.06 | -0.01 | -0.03 | 0.02 | 0.03 | 0.03 | 0.03 |
| | (0.07) | (0.06) | (0.07) | (0.05) | (0.03) | (0.03) | (0.03) | (0.03) |
| LC & SG | 0.07 | 0.04 | 0.11* | 0.13** | 0.05 | 0.07 | 0.11** | 0.09** |
| | (0.06) | (0.04) | (0.06) | (0.05) | (0.09) | (0.09) | (0.05) | (0.04) |
| SG | -0.01 | 0.01 | -0.02 | 0.04 | -0.02 | -0.05 | -0.01 | 0.01 |
| | (0.05) | (0.04) | (0.06) | (0.05) | (0.08) | (0.10) | (0.05) | (0.04) |
| LC | -0.01 | -0.01 | 0.03 | 0.03 | -0.03 | -0.06 | 0.03 | 0.02 |
| | (0.05) | (0.04) | (0.06) | (0.05) | (0.09) | (0.10) | (0.05) | (0.04) |
| Mean Control group | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 |
| R-squared | 0.416 | 0.580 | 0.415 | 0.580 | 0.415 | 0.580 | 0.415 | 0.580 |
| Observations | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 |
| No. of clusters | 200 | 200 | 200 | 200 | 200 | 200 | 200 | 200 |

Notes: LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both Learning Camps and Study Groups), to the Study Groups alone, or to the Learning Camps alone, respectively. All regressions include the variable being interacted as well (coefficient omitted). *Basic controls* include age and gender of the child, baseline value of the dependent variable and grade fixed effects. Standard errors clustered by village in parentheses. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

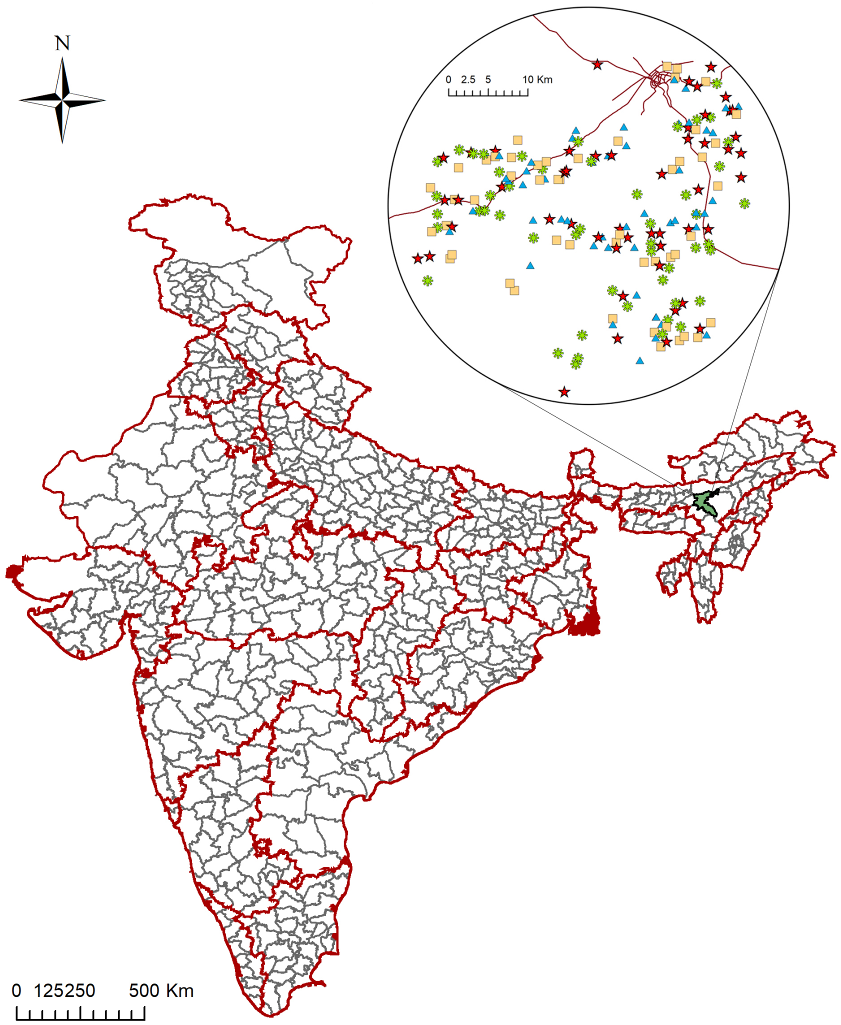
**Table 4:** Program impact on children, caregivers, and schools

| **Panel A: Child outcomes** | Attendance | Study outside | Participate in SG | Study up to | School satisf. (PCA) | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | |
| LC & SG | 0.01 | 0.07*** | 0.21*** | -0.06 | 0.00 | |
| | (0.03) | (0.02) | (0.03) | (0.19) | (0.15) | |
| SG | -0.01 | 0.05** | 0.14*** | -0.18 | -0.14 | |
| | (0.04) | (0.02) | (0.03) | (0.16) | (0.16) | |
| LC | 0.01 | 0.04 | 0.01 | -0.25 | 0.01 | |
| | (0.04) | (0.02) | (0.01) | (0.19) | (0.15) | |
| Mean Control group | 0.75 | 0.77 | 0.03 | 11.26 | 0.04 | |
| Observations | 5,328 | 5,328 | 5,328 | 4,329 | 5,108 | |
| p-val(LC & SG=SG) | 0.61 | 0.45 | 0.05 | 0.53 | 0.37 | |
| p-val(LC & SG=LC) | 0.97 | 0.18 | 0.00 | 0.35 | 0.97 | |
| **Panel B: Household inputs** | Involved in SG | Time w. child (%) | Help w. homework | Education spend. | School interac. (PCA) | ASTE |
| | (1) | (2) | (3) | (4) | (5) | (1)-(5) |
| LC & SG | 0.03*** | -0.00 | 0.12 | 0.06 | 0.21** | 0.13*** |
| | (0.01) | (0.01) | (0.23) | (0.32) | (0.09) | (0.04) |
| SG | 0.02*** | -0.01 | 0.17 | 0.05 | 0.19** | 0.10** |
| | (0.01) | (0.01) | (0.29) | (0.34) | (0.09) | (0.04) |
| LC | 0.00 | 0.00 | 0.20 | -0.16 | 0.20** | 0.06* |
| | (0.01) | (0.01) | (0.24) | (0.32) | (0.08) | (0.03) |
| Mean Control group | 0.00 | 0.24 | 5.29 | 3.18 | -0.15 | |
| Observations | 4,265 | 4,251 | 4,250 | 4,265 | 4,246 | 4,265 |
| p-val(LC & SG=SG) | 0.33 | 0.78 | 0.84 | 0.98 | 0.84 | |
| p-val(LC & SG=LC) | 0.00 | 0.64 | 0.68 | 0.31 | 0.92 | |
| **Panel C: School inputs** | Teachers pres. (%) | Material invest. (PCA) | Interac. w. SMC | Interac. w. PTA | Interac. w. MoE | ASTE |
| | (1) | (2) | (3) | (4) | (5) | (1)-(5) |
| LC & SG | 0.01 | -0.40* | -0.20** | -0.17 | -0.15** | -0.34*** |
| | (0.05) | (0.23) | (0.09) | (0.11) | (0.07) | (0.11) |
| SG | 0.02 | -0.47** | -0.04 | 0.04 | -0.11 | -0.18* |
| | (0.05) | (0.23) | (0.08) | (0.10) | (0.08) | (0.10) |
| LC | 0.06 | -0.36 | -0.22** | -0.12 | -0.14* | -0.28** |
| | (0.04) | (0.24) | (0.09) | (0.10) | (0.07) | (0.11) |
| Mean Control group | 0.79 | 0.30 | 0.84 | 0.56 | 0.90 | |
| Observations | 199 | 197 | 199 | 186 | 197 | 199 |
| p-val(LC & SG=SG) | 0.76 | 0.78 | 0.09 | 0.05 | 0.67 | |
| p-val(LC & SG=LC) | 0.22 | 0.82 | 0.83 | 0.64 | 0.89 | |

Notes: In Panel A, the dependent variables are defined as follows: child at endline was present in school on survey day (1); child studies outside school (2); child has participated in a study group (3); child would like to study until grade X (4); principal component analysis (PCA) across 19 variables on child's school satisfaction and learning support (5). In Panel B, the dependent variables are defined as follows: respondent is involved in an after-school study group (1); share of time in an average day spent studying, playing, or reading with the child (2); hours per week spent helping child with homework (3); total amount spent on education-related items for the child (4); principal component analysis (PCA) across 5 variables on the caregiver's interaction with the school (5); ASTE of columns 1-5 (6). The varying number of observations is due to some respondents not knowing the answer to some specific questions. In Panel C, the dependent variables are defined as follows: share of teachers present at the school on the day of the survey (1); principal component analysis (PCA) over any investments undertaken in school in the previous 14 months in: (i) construction, (ii) repair, (iii) purchase of teaching-learning material, and (iv) purchase of sitting mats for children (2); any School Management Committee meeting in the previous 2 months (3); any parent-teacher meeting in the previous 2 months (4); any visit from the Ministry of Education in the previous 3 months (5); ASTE of columns 1-5 (6). See Table **??** for a list of *basic controls* included in Panels A and B. Baseline values not included for Participation in SG, Help with homework, and PCA and ASTE regressions. LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both Learning Camps and Study Groups), to the Study Groups alone, or to the Learning Camps alone, respectively. The p-values in the bottom of the table are the test of the null hypothesis of equal treatment effects between different intervention arms. Standard errors clustered by village in parentheses. Randomization inference-based p-values based on 1,000 iterations reported in square brackets. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

# Appendix A – Additional Figures and Tables

**Figure A.1:** Study Locations




Note: The study was set in Nagaon district, within the State of Assam. We drew a sample of 200 villages from the list of villages that Pratham deemed eligible to receive the program. The villages were then randomly assigned to one of the four study arms (50 villages per arm). The map above illustrates the location of Nagaon district within India, as well as an enlarged view on the location of the 200 study villages. The four different symbols indicate assignment to the different study arms.

**Figure A.2:** Timeline

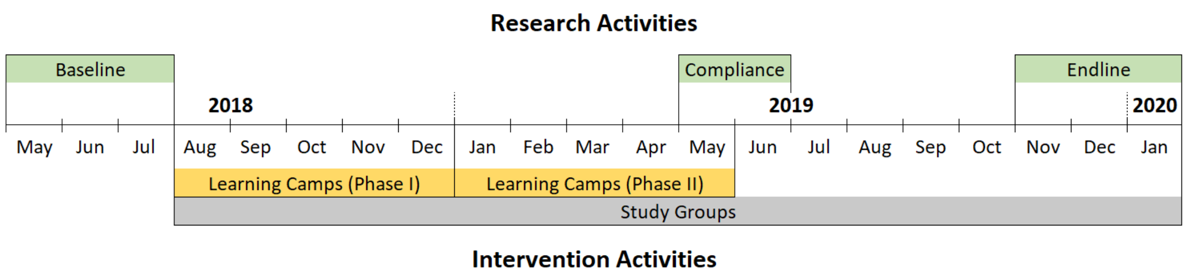



**Figure A.3:** Distribution of ASER math levels by grade at baseline

**Figure A.4:** Distribution of ASER language levels by grade at baseline

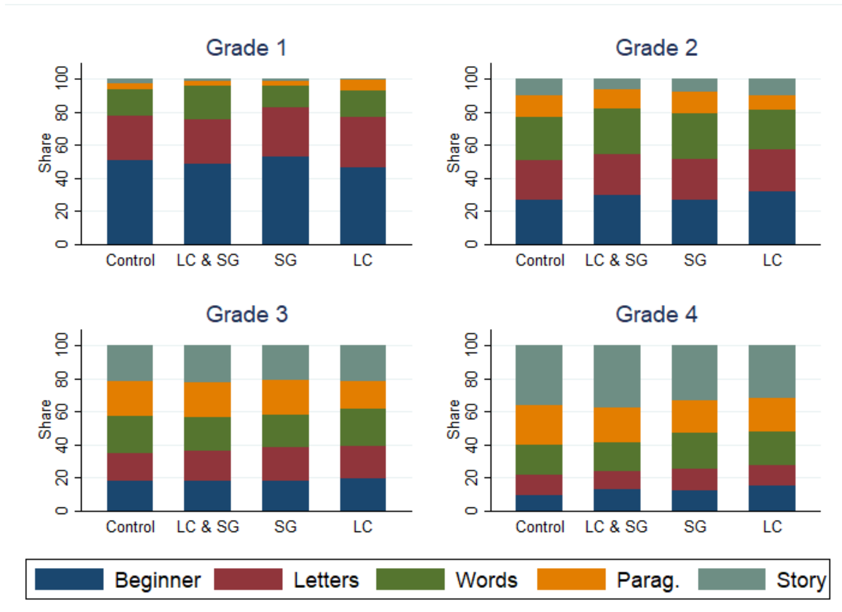
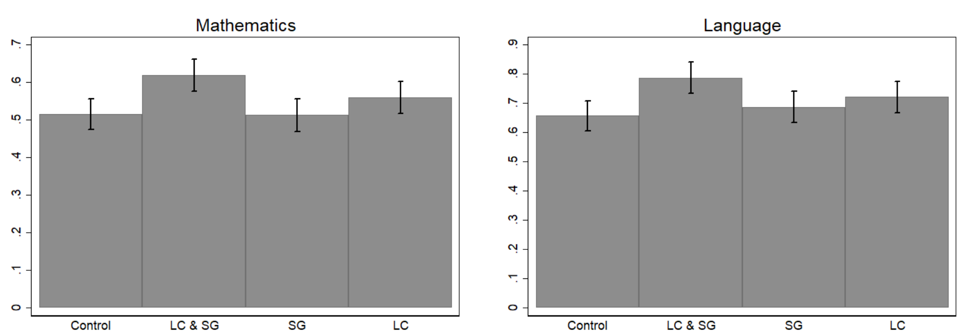



**Figure A.5:** Learning gains in mathematics and language over the study period




Notes: The figures show the difference in learning level in mathematics (left) and language (right), between endline and baseline, across the entire sample, by treatment arm.

**Figure A.6:** Distribution of ASER math levels by grade at endline

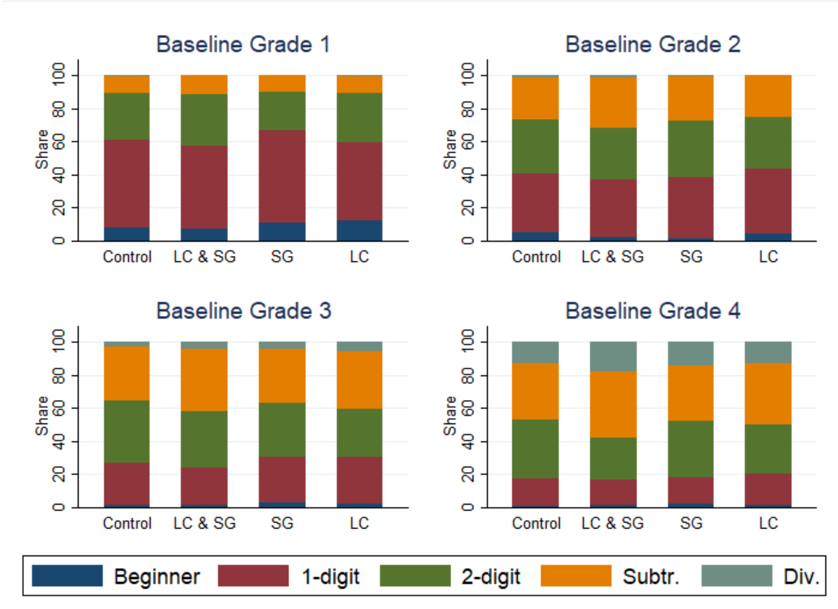



**Figure A.7:** Distribution of ASER language levels by grade at endline

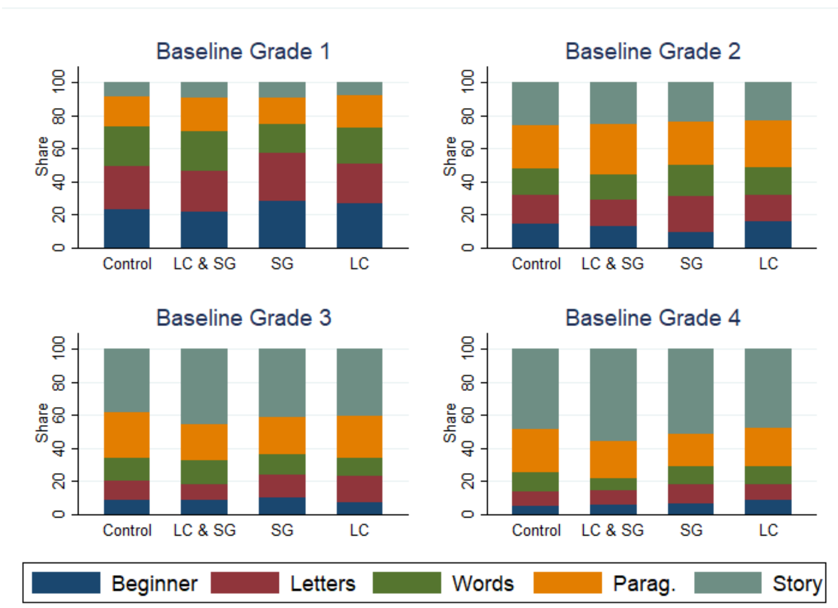
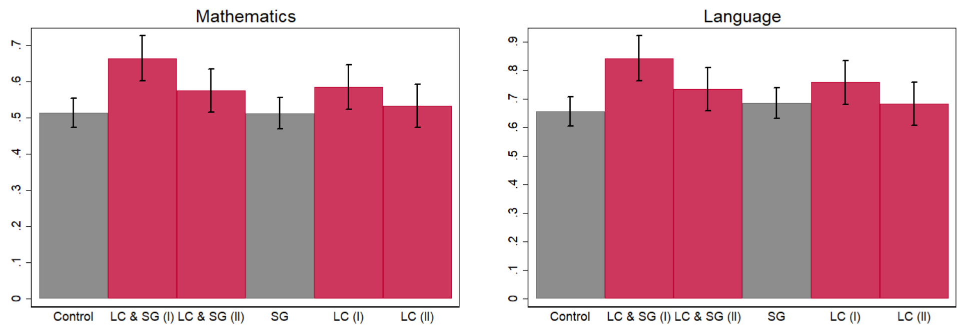


**Figure A.8:** Cost effectiveness of other education programs, from RCTs

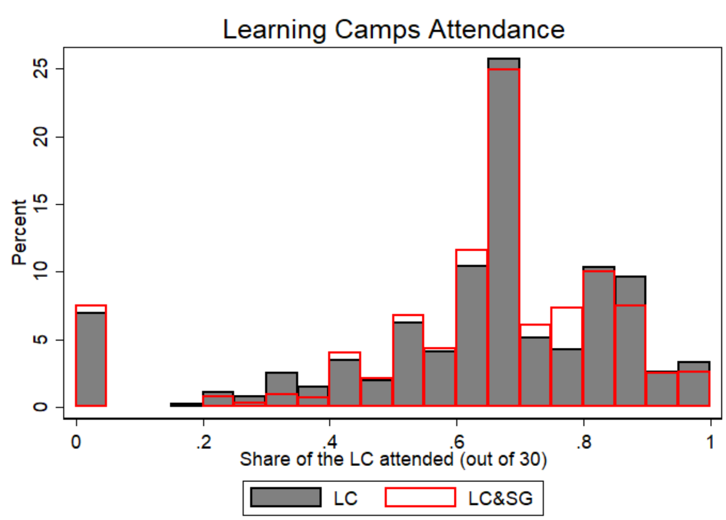

Source: Bhula et al (2013)

**Figure A.9:** Learning gains in mathematics and language over the study period, by phase




Notes: The figures show the difference in learning level in mathematics (left) and language (right), between endline and baseline, across the entire sample, by treatment arm. Compared to Figure A.5, this figure distinguishes (in red) between villages that received the learning camps in the first implementation phase (I) and those that received them in the second one (II).

**Figure A.10:** Learning Camps Attendance




Notes: The figure show the distribution of the share of Learning Camps days attended by the students in our sample, across the Learning Camps only study arm (gray bars) and the full program study arm (red bars). The total number of Learning Camps days was 30.

**Table A.1:** Comparison of sample to schools in Nagaon and Assam

| | (1) Assam state | | (2) Nagaon district | | (3) Sample | | T-test Difference | |
|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean/SE | N | Mean/SE | N | Mean/SE | (3)-(1) | (3)-(2) |
| Urban (Y/N) | 57732 | 0.036 | 3112 | 0.054 | 200 | 0.025 | -0.011 | -0.029* |
| | | (0.001) | | (0.004) | | (0.011) | | |
| No. classrooms | 57732 | 2.314 | 3112 | 2.459 | 200 | 2.685 | 0.371*** | 0.226 |
| | | (0.008) | | (0.044) | | (0.087) | | |
| No. students | 57732 | 65.518 | 3112 | 127.949 | 200 | 54.425 | -11.093** | -73.524*** |
| | | (0.260) | | (2.381) | | (1.530) | | |
| Teach-to-stud. ratio | 57158 | 0.098 | 3093 | 0.077 | 200 | 0.089 | -0.009 | 0.012* |
| | | (0.001) | | (0.002) | | (0.012) | | |
| Share teach. w/ prof. qual. | 57428 | 0.635 | 3103 | 0.597 | 200 | 0.730 | 0.095*** | 0.133*** |
| | | (0.002) | | (0.008) | | (0.025) | | |
| Instr. in Assamese (Y/N) | 57732 | 0.735 | 3112 | 0.972 | 200 | 0.995 | 0.260*** | 0.023* |
| | | (0.002) | | (0.003) | | (0.005) | | |
| School funds received | 57732 | 7910.585 | 3112 | 7216.355 | 200 | 7460.000 | -450.585 | 243.645 |
| | | (145.476) | | (350.402) | | (294.918) | | |
| Electricity (Y/N) | 57732 | 0.155 | 3112 | 0.337 | 200 | 0.190 | 0.035 | -0.147*** |
| | | (0.002) | | (0.008) | | (0.028) | | |
| Playground (Y/N) | 57732 | 0.556 | 3112 | 0.691 | 200 | 0.585 | 0.029 | -0.106*** |
| | | (0.002) | | (0.008) | | (0.035) | | |
| CCE implemented (Y/N) | 57732 | 0.912 | 3112 | 0.905 | 200 | 0.920 | 0.008 | 0.015 |
| | | (0.001) | | (0.005) | | (0.019) | | |
| Pupil records maintained (Y/N) | 57732 | 0.887 | 3112 | 0.863 | 200 | 0.880 | -0.007 | 0.017 |
| | | (0.001) | | (0.006) | | (0.023) | | |
| SMC instituted (Y/N) | 57732 | 0.936 | 3112 | 0.866 | 200 | 1.000 | 0.064*** | 0.134*** |
| | | (0.001) | | (0.006) | | (0.000) | | |
| Workdays | 57732 | 192.920 | 3112 | 171.208 | 200 | 244.130 | 51.210*** | 72.922*** |
| | | (0.373) | | (1.990) | | (1.269) | | |
| Teach. work. hrs/day | 57732 | 4.903 | 3112 | 4.452 | 200 | 6.374 | 1.471*** | 1.922*** |
| | | (0.010) | | (0.053) | | (0.078) | | |
| No. of inspections | 57732 | 2.435 | 3112 | 2.553 | 200 | 3.810 | 1.375*** | 1.257*** |
| | | (0.020) | | (0.091) | | (0.441) | | |

Notes: Data comes from the 2017-18 Unified District Information System for Education (UDISE) kept by the Ministry of Education. CCE indicates Continuous and Comprehensive Evaluation assessment process. The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

**Table A.2:** Pre-intervention characteristics of schools, children, and caregivers

| | C | SG & LC | SG | LC | p-value |
|---|---|---|---|---|---|
| **Panel A: Schools (N= 200)** | | | | | |
| Number of classrooms | 2.26 | 2.00 | 2.06 | 2.14 | 0.79 |
| | (1.12) | (1.73) | (1.39) | (1.11) | |
| Total teaching staff | 4.12 | 3.70 | 3.58 | 3.98 | 0.64 |
| | (2.56) | (3.29) | (2.07) | (2.14) | |
| Total enrollment | 55.72 | 52.52 | 52.92 | 52.60 | 0.83 |
| | (19.85) | (20.57) | (23.42) | (20.73) | |
| Share of girls enrolled | 0.51 | 0.49 | 0.51 | 0.50 | 0.68 |
| | (0.11) | (0.07) | (0.08) | (0.07) | |
| **Panel B: Children (N= 5726)** | | | | | |
| Age | 7.69 | 7.64 | 7.61 | 7.61 | 0.77 |
| | (1.59) | (1.55) | (1.53) | (1.56) | |
| Girl | 0.51 | 0.49 | 0.50 | 0.50 | 0.80 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| Present in school | 0.68 | 0.75 | 0.74 | 0.68 | 0.35 |
| | (0.47) | (0.44) | (0.44) | (0.47) | |
| Likes going to school [1-5] | 4.51 | 4.54 | 4.57 | 4.59 | 0.33 |
| | (0.89) | (0.85) | (0.81) | (0.78) | |
| Study outside school | 0.71 | 0.75 | 0.74 | 0.76 | 0.25 |
| | (0.45) | (0.43) | (0.44) | (0.43) | |
| ASER score (language) | 1.75 | 1.72 | 1.66 | 1.63 | 0.78 |
| | (1.44) | (1.43) | (1.42) | (1.41) | |
| ASER score (math) | 1.42 | 1.43 | 1.40 | 1.35 | 0.67 |
| | (0.84) | (0.85) | (0.89) | (0.85) | |
| Tracked at endline (share) | 0.94 | 0.93 | 0.92 | 0.93 | 0.59 |
| | (0.24) | (0.25) | (0.28) | (0.25) | |
| **Panel C: Primary Caregivers (N= 4592)** | | | | | |
| # HH members | 5.24 | 5.15 | 5.21 | 5.30 | 0.55 |
| | (1.76) | (1.65) | (1.70) | (1.82) | |
| # children enrolled | 1.44 | 1.42 | 1.47 | 1.43 | 0.61 |
| | (0.82) | (0.78) | (0.82) | (0.84) | |
| Asset Index | -0.10 | 0.19 | -0.13 | 0.04 | 0.23 |
| | (1.89) | (1.78) | (1.81) | (1.83) | |
| Primary caregiver is literate | 0.48 | 0.53 | 0.51 | 0.50 | 0.73 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| Would like child to go to university | 0.44 | 0.45 | 0.45 | 0.49 | 0.61 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| Pays tuitions | 0.19 | 0.24 | 0.19 | 0.21 | 0.27 |
| | (0.39) | (0.43) | (0.40) | (0.41) | |
| Estimates language level correctly | 0.33 | 0.34 | 0.30 | 0.33 | 0.37 |
| | (0.47) | (0.47) | (0.46) | (0.47) | |
| Overestimates language level | 0.50 | 0.50 | 0.51 | 0.52 | 0.87 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| Estimates math level correctly | 0.39 | 0.37 | 0.38 | 0.36 | 0.69 |
| | (0.49) | (0.48) | (0.49) | (0.48) | |
| Overestimates math level | 0.52 | 0.54 | 0.52 | 0.56 | 0.49 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| Tracked at endline (share) | 0.95 | 0.93 | 0.93 | 0.94 | 0.38 |
| | (0.22) | (0.25) | (0.26) | (0.24) | |

Notes: Standard deviations shown in parentheses. C, LC & SG, SG and LC indicate that the observation belongs to one of the following four study groups: control, full program (which included both Learning Camps and Study Groups), Study Groups alone, or Learning Camps alone. The last column shows the p-value from testing whether the mean is equal across all treatment groups ($H_0$ := mean is equal across groups).

**Table A.3:** Attrition by grade

| | Child survey | | | | Caregiver survey | | | |
|---|---|---|---|---|---|---|---|---|
| Grade at baseline | Baseline | Endline | Attrition | Attrition (%) | Baseline | Endline | Attrition | Attrition (%) |
| Grade 1 | 1449 | 1334 | 115 | 0.079 | 1168 | 1091 | 77 | 0.066 |
| Grade 2 | 1373 | 1277 | 96 | 0.070 | 1106 | 1032 | 74 | 0.067 |
| Grade 3 | 1460 | 1349 | 111 | 0.076 | 1160 | 1086 | 74 | 0.064 |
| Grade 4 | 1444 | 1368 | 76 | 0.053 | 1158 | 1094 | 64 | 0.055 |

**Table A.4:** Attrition checks

| *Dep Var:* | Attrited Child | | | | | |
|---|---|---|---|---|---|---|
| *Interaction with...:* | | | Grade | Girl | ASER | |
| | | | | | Math | Lang. |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| LC & SG | 0.004 | 0.006 | 0.016 | 0.000 | 0.002 | -0.001 |
| | (0.014) | (0.014) | (0.028) | (0.018) | (0.027) | (0.021) |
| SG | 0.019 | 0.021 | 0.038 | 0.018 | 0.025 | 0.033 |
| | (0.015) | (0.015) | (0.028) | (0.018) | (0.027) | (0.024) |
| LC | 0.007 | 0.009 | -0.005 | 0.012 | -0.005 | -0.002 |
| | (0.015) | (0.014) | (0.025) | (0.019) | (0.024) | (0.020) |
| LC & SG × ... | | | -0.004 | 0.013 | 0.004 | 0.004 |
| | | | (0.009) | (0.019) | (0.013) | (0.006) |
| SG × ... | | | -0.007 | 0.007 | -0.003 | -0.008 |
| | | | (0.009) | (0.017) | (0.012) | (0.008) |
| LC × ... | | | 0.005 | -0.007 | 0.009 | 0.006 |
| | | | (0.008) | (0.018) | (0.011) | (0.006) |
| Basic controls | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean Control group | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| R-squared | 0.001 | 0.017 | 0.017 | 0.017 | 0.022 | 0.025 |
| Observations | 5,726 | 5,726 | 5,726 | 5,726 | 5,726 | 5,726 |
| No. of clusters | 200 | 200 | 200 | 200 | 200 | 200 |

Notes: LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the standard Pratham program (which included both learning camps and community Study Groups), to the commmunity study groups alone, or to the learning camps alone, respectively. All regressions with interactions include interaction components as well (not reported). *Basic controls* include: grade fixed effects, gender, and age of the child. Standard errors clustered by village in parentheses. *Significant at 10% level; **Significant at 5% level;, ***Significant at 1% level.

**Table A.5:** Self-reported program exposure

| **Panel A: Children** | Interaction with Pratham | | Learning Camps | | | Study Groups | |
|---|---|---|---|---|---|---|---|
| | Heard of | Interacted with | Pratham Tested | Pratham Teachers | Divided in groups | in village | participated |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| LC & SG | 0.16*** | 0.36*** | 0.13*** | 0.10*** | 0.35*** | 0.24*** | 0.21*** |
| | (0.03) | (0.03) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| SG | 0.05** | 0.21*** | 0.03** | 0.01 | 0.16*** | 0.16*** | 0.14*** |
| | (0.02) | (0.03) | (0.02) | (0.01) | (0.03) | (0.03) | (0.03) |
| LC | 0.10*** | 0.27*** | 0.08*** | 0.06*** | 0.28*** | 0.01 | 0.01 |
| | (0.02) | (0.03) | (0.01) | (0.01) | (0.03) | (0.01) | (0.01) |
| Mean Control group | 0.08 | 0.34 | 0.05 | 0.03 | 0.24 | 0.07 | 0.03 |
| Observations | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 |

| **Panel B: Caregivers** | Pratham | | | Learning Camps | | | Study |
|---|---|---|---|---|---|---|---|
| | Heard of | Interacted | Material | In school | TL activities | Diff. gr. | Groups |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| LC & SG | 0.15*** | 0.09*** | 0.36*** | 0.12*** | 0.10*** | 0.23*** | 0.26*** |
| | (0.02) | (0.02) | (0.03) | (0.02) | (0.01) | (0.03) | (0.03) |
| SG | 0.09*** | 0.06*** | 0.25*** | 0.08*** | 0.06*** | 0.14*** | 0.23*** |
| | (0.02) | (0.01) | (0.03) | (0.02) | (0.01) | (0.03) | (0.03) |
| LC | 0.08*** | 0.03** | 0.24*** | 0.05*** | 0.04*** | 0.13*** | 0.04*** |
| | (0.02) | (0.01) | (0.02) | (0.01) | (0.01) | (0.02) | (0.01) |
| Mean Control group | 0.07 | 0.02 | 0.10 | 0.02 | 0.02 | 0.05 | 0.02 |
| Observations | 4,251 | 4,224 | 4,265 | 4,251 | 4,251 | 4,265 | 4,265 |

Notes: The dependent variables in Panel A are taken from the children's surveys and are indicators capturing whether: child ever heard of Pratham (1); indicator capturing any interaction with Pratham (2); Pratham performed any test to assess learning levels (3); Pratham staff replaced the regular teacher (4); students were divided in groups, based on learning level, and combined with students from other grades (5); Study Groups are arranged in the village after school (6); child participated in these groups (7). The dependent variables in Panel B are taken from the caregivers' surveys and are indicators capturing whether: respondent ever heard of Pratham (1); respondent ever interacted with Pratham (2); respondent ever saw Pratham's teaching-learning material (3); Pratham worked in the school (4); Pratham placed the child in a different grade than his/her usual (5); the child worked in groups different than his/her usual class (6); any Study Groups was arranged in the village (7). LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both learning camps and community Study Groups), to the study groups alone, or to the learning camps alone, respectively. All regressions include students' gender, age, and grade fixed effects. The varying number of observations in Panel B is due to some respondents not knowing the answer to some specific questions. Standard errors clustered by village in parentheses. Randomization inference-based p-values based on 1,000 iterations reported in square brackets. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

**Table A.6:** School-reported program exposure

| *Dep Var:* | Pratham | | Learning Camps | | Study |
|---|---|---|---|---|---|
| | Heard of | Interacted with | Direct question | Students in groups | Groups in village |
| | (1) | (2) | (3) | (4) | (5) |
| LC & SG | 0.58*** | 0.92*** | 0.22*** | 0.62*** | 0.74*** |
| | (0.07) | (0.04) | (0.06) | (0.08) | (0.06) |
| SG | 0.46*** | 0.78*** | 0.10** | 0.23** | 0.55*** |
| | (0.09) | (0.06) | (0.05) | (0.09) | (0.07) |
| LC | 0.54*** | 0.88*** | 0.22*** | 0.58*** | 0.34*** |
| | (0.08) | (0.05) | (0.06) | (0.08) | (0.07) |
| Mean Control | 0.38 | 0.02 | 0.02 | 0.16 | 0.02 |
| R-squared | 0.309 | 0.634 | 0.064 | 0.264 | 0.310 |
| Observations | 199 | 199 | 199 | 199 | 199 |
| p-val(LC & SG=SG) | 0.04 | 0.03 | 0.13 | 0.00 | 0.05 |
| p-val(LC & SG=LC) | 0.40 | 0.47 | 1.00 | 0.64 | 0.00 |

Notes: LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both learning camps and community Study Groups), to the study groups alone, or to the learning camps alone, respectively. Robust standard errors in parenthesis. Randomization inference p-values in square bracket. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

**Table A.7:** Program impact using alternative learning outcomes measures

| *Dep Var:* | Alternative Learning Outcomes | | | | | |
|---|---|---|---|---|---|---|
| *Subject:* | Math | Lang. | Full | Math | Lang. | Full |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| LC & SG | 0.10** | 0.04 | 0.09*** | | | |
| | (0.05) | (0.04) | (0.03) | | | |
| SG | 0.01 | -0.05 | -0.00 | 0.01 | -0.05 | -0.00 |
| | (0.05) | (0.04) | (0.04) | (0.05) | (0.04) | (0.04) |
| LC | 0.02 | 0.01 | 0.03 | | | |
| | (0.05) | (0.04) | (0.03) | | | |
| LCI & SG | | | | 0.15*** | 0.07 | 0.12*** |
| | | | | (0.06) | (0.05) | (0.04) |
| LCII & SG | | | | 0.05 | 0.02 | 0.05 |
| | | | | (0.06) | (0.05) | (0.04) |
| LC Phase I | | | | 0.04 | 0.01 | 0.05 |
| | | | | (0.06) | (0.04) | (0.04) |
| LC Phase II | | | | -0.01 | 0.01 | 0.01 |
| | | | | (0.06) | (0.05) | (0.04) |
| Basic controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean Control group | 0.00 | 0.00 | -0.01 | 0.00 | 0.00 | -0.01 |
| R-squared | 0.459 | 0.530 | 0.677 | 0.460 | 0.530 | 0.678 |
| Observations | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 | 5,328 |
| No. of clusters | 200 | 200 | 200 | 200 | 200 | 200 |

Notes: The dependent variables are alternative measures of learning outcomes: test scores in the mathematics section of "Test A" (normalized using mean and standard deviation of the control group); test scores in the lnguage section of "Test A" (normalized using mean and standard deviation of the control group); comprehensive test score obtained from a two parameter logistic (2PL) item response theory (IRT) model, combing all ASER and Test A questions. LC & SG, SG and LC are indicator variables that take on value 1 if the child's village was assigned to the full program (which included both learning camps and community Study Groups), to the study groups alone, or to the learning camps alone, respectively. *Basic controls* include age and gender of the child, baseline value of the dependent variable and grade fixed effects. Standard errors clustered by village in parentheses.*Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

**Table A.8:** Comparison of the *Learning Camps* vs Banerjee et al. (2016)

| | **This study** | **Banerjee et al (2016)** |
|---|---|---|
| Setting | Assam | Uttar Pradesh |
| Year | 2018-2019 | 2013-2014 |
| Implementer | Pratham | Pratham |
| Data Collection | J-PAL | J-PAL |
| Assessment | ASER | ASER |
| Study length | $\sim$18 months | $\sim$12 months |
| LC Duration | 32 (30+2) days | 50 (40+10) days |
| Grades covered | 1 to 5 | 3 to 5 |
| Enrollment /school (3-5) | 32 | 72 |
| ASER math (avg) | 2.25 | 2.11 |
| ASER language (avg) | 2.84 | 1.97 |
| ATE | 0.0 SD | 0.6-0.7 SD |

Notes: *Enrollment /school* indicates the average number of students enrolled in grades 3 to 5 in the schools included in the sample. *ASER math* and *ASER language* indicate the average level in math and language, respectively, from comparable tests modeled around the ASER test. To ensure comparability, average test scores only includes assessments performed at the end of the school year (i.e. endline for both studies) and considers students enrolled in grades 3 to 5 in the control group schools.

**Table A.9:** *Learning Camps* in Assam vs Uttar Pradesh

| *Dep Var:* | ASER score (SD) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Math | Lang. | Math | Lang. | Math | Lang. | Math | Lang. |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| LC × Assam | -0.584*** | -0.691*** | -0.564*** | -0.628*** | -0.701*** | -0.709*** | | |
| | (0.089) | (0.092) | (0.052) | (0.048) | (0.068) | (0.062) | | |
| LC | 0.596*** | 0.659*** | 0.614*** | 0.666*** | 0.470*** | 0.680*** | 0.459*** | 0.819*** |
| | (0.040) | (0.038) | (0.024) | (0.023) | (0.168) | (0.150) | (0.173) | (0.159) |
| Assam | 0.065 | 0.075 | -0.200*** | -0.109*** | -0.134*** | -0.074** | | |
| | (0.058) | (0.059) | (0.042) | (0.037) | (0.042) | (0.038) | | |
| ASER Math, BL | | | 0.645*** | | 0.647*** | | | |
| | | | (0.010) | | (0.012) | | | |
| ASER Lang, BL | | | | 0.740*** | | 0.799*** | | |
| | | | | (0.009) | | (0.011) | | |
| Girl | | | -0.089*** | 0.006 | -0.099*** | -0.029* | -0.131*** | -0.059*** |
| | | | (0.016) | (0.014) | (0.019) | (0.016) | (0.021) | (0.017) |
| Age | | | -0.004 | -0.027*** | 0.010 | -0.011 | 0.015 | -0.011 |
| | | | (0.009) | (0.008) | (0.011) | (0.010) | (0.012) | (0.011) |
| School size | | | -0.003*** | -0.002*** | -0.002*** | -0.001*** | -0.002*** | -0.001*** |
| | | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Study Months | | | 0.010** | 0.005 | 0.004 | -0.001 | 0.002 | -0.001 |
| | | | (0.005) | (0.004) | (0.003) | (0.003) | (0.003) | (0.003) |
| Grade 4 | | | 0.121*** | 0.072*** | 0.072*** | 0.052** | 0.067** | 0.075*** |
| | | | (0.019) | (0.017) | (0.024) | (0.021) | (0.027) | (0.023) |
| LC × Girl | | | | | 0.018 | 0.069** | 0.042 | 0.079*** |
| | | | | | (0.032) | (0.028) | (0.034) | (0.030) |
| LC × Age | | | | | -0.030* | -0.032* | -0.032* | -0.031* |
| | | | | | (0.017) | (0.017) | (0.019) | (0.018) |
| LC × School size | | | | | -0.002*** | -0.001 | -0.002*** | -0.001 |
| | | | | | (0.001) | (0.001) | (0.001) | (0.001) |
| LC × Study month | | | | | 0.013 | 0.011* | 0.014* | 0.011* |
| | | | | | (0.008) | (0.007) | (0.008) | (0.007) |
| LC × Grade 4 | | | | | 0.105*** | 0.039 | 0.113*** | 0.031 |
| | | | | | (0.038) | (0.033) | (0.041) | (0.036) |
| LC × ASER Math (SD) | | | | | -0.005 | | | |
| | | | | | (0.020) | | | |
| LC × ASER Lang (SD) | | | | | | -0.130*** | | |
| | | | | | | (0.017) | | |
| LC × ASER Math | | | | | | | -0.014 | |
| | | | | | | | (0.021) | |
| LC × ASER Lang | | | | | | | | -0.122*** |
| | | | | | | | | (0.015) |
| Sample | Assam, UP | Assam, UP | Assam, UP | Assam, UP | Assam, UP | Assam, UP | UP | UP |
| R-squared | 0.069 | 0.094 | 0.402 | 0.525 | 0.403 | 0.528 | 0.407 | 0.516 |
| Observations | 13,839 | 13,837 | 13,601 | 13,585 | 13,601 | 13,585 | 11,572 | 11,556 |

Notes: *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

# Appendix B – Test Administration

The tests were administered orally and individually to each child by trained enumerators, in the local (Assamese) language. The content was based on well-established assessment tools that have already been extensively piloted and used across India. As mentioned in the main text, two tests were originally used in this project.

The first test mirrored the standard ASER (Annual Status of Education Report) test, a nationwide test conducted yearly by the ASER Center all over India for children aged 5 to 16. The test is divided in a math and a language component and had already been extensively piloted, tested, and used in previous evaluations conducted in India (e.g. Banerjee et al., 2016; Banerji et al., 2017). According to the ASER test, children are classified in five categories based on their knowledge. For mathematics the categories are:

- Beginner (no number recognition)
- single-digit number recognition
- double-digit number recognition
- subtraction (of double-digit numbers)
- division (of a double-digit number by a single digit)

For language the categories are:

- Beginner (cannot recognize letters)
- recognize letters
- reads words
- reads a paragraph

- reads a short story

Figure B.1 shows an English version of a standard ASER test.

The second “test A”, was instead created by the research team, based on tests that had been previously used in other studies conducted in India (Muralidharan and Singh, 2020). Figure B.2 shows a sample of questions included in the test. By endline, this alternative test resulted too easy for the students in the sample and the distribution of marks is severely skewed (on average, across the full sample students completed correctly 64% of the mathematics section and 76% of the language section). For this reason, in our main analysis we excluded this test, although results for the mathematics section (where the problem was less severe) are very consistent. To further check the robustness of our findings, we also combine all our tests (ASER and Alternative Test) in a single measure, using a two parameter logistic (2PL) item response theory (IRT) model (see for instance Jacob and Rothstein (2016)). Results are reported in Table A.7 in Appendix A.

**Figure B.1:** ASER Test – sample questions


READING TEST SAMPLE (1)

Letter

b s o
k m
y r h
t x

Ask the child to recognize any 5 letters. At least 4 must be correct.

Word

ring bad
ball
cold king
clap foot
fan
girl crow

Ask the child to read any 5 words. At least 4 must be correct.

READING TEST SAMPLE (1)

Story

Salma is a little girl. She had a pretty doll. She loved playing with her doll. One day the doll fell from her hand to the floor. It broke into many pieces. Salma was very sad. She cried a lot. Her mother gave her another doll. Now she is happy again.

Para

Ravi is a boy.
He has many friends.
He loves to draw.
He does not like to sing.

Para

My village is very big.
It has many houses.
It also has a shop.
The bus stops in my village.

MATH TEST SAMPLE (1)

| Number recognition 1—9 | | Number recognition 10—99 | | Subtraction | | Division |
|---|---|---|---|---|---|---|
| 1 | 4 | 51 | 83 | 46 − 29 | 63 − 39 | 7) 879 |
| 7 | 3 | 37 | 65 | 47 − 28 | 45 − 17 | 6) 824 |
| 6 | 9 | 55 | 26 | 92 − 76 | 84 − 57 | 8) 985 |
| 5 | 2 | 91 | 43 | 52 − 14 | 66 − 48 | 4) 517 |
| | | 36 | 27 | | | |
| Ask the child to recognize any 5 numbers. At least 4 must be correct. | | Ask the child to recognize any 5 numbers. At least 4 must be correct. | | Ask the child to do any 2 subtraction problems. Both must be correct. | | Ask the child to do any 1 division problem. It must be correct. |

**Figure B.2:** Test A – sample questions