Economics of Education Review 96 (2023) 102439

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

In good times and bad: Low-cost mobile teaching during a pandemic

Manini Ojha[a], Kartik Yadav[b,*]

[a] Jindal School of Government and Public Policy, OP Jindal Global University, Sonipat, India
[b] Indian Institute of Management Nagpur, MIHAN, Nagpur, India

A R T I C L E I N F O

JEL classification:
I2
I21
I24
I25
O15

Keywords:
COVID-19
Learning outcomes
Schooling
Home-based learning
Mobile
India

A B S T R A C T

In view of school closures due to the COVID-19 pandemic, this paper examines how a Home-Based Learning program affects learning outcomes of children in under-resourced communities. To overcome limited internet connectivity, the program provides remote instructions via phone calls and simple text messages along with automated voice calls to engage children enrolled in grades one to five in activity-based learning content. This intervention was conducted in three districts in the state of Odisha in India. Using a difference-in-differences framework, we find that the intervention led to a statistically significant improvement in basic number recognition and arithmetic operations, and language learning scores of children by 4.69 percentage points and 5.52 percentage points, respectively. Our results are robust to alternative methods of estimation and application of Lee bounds, thus indicating that well-designed low cost interventions could be a useful supplement for continued learning in the face of sudden shocks in low income countries. With a rise in hybrid format of teaching and learning, such interventions have the capability to cushion the decline in learning levels and provide a safety net in the event of school closures.

1. Introduction

The COVID-19 pandemic paralyzed education systems across the globe and created a devastating educational emergency. Partial or full school closures owing to the pandemic affected almost 168 million children globally and it was expected that about a 100 million additional children would fall below the minimum proficiency level in reading as a result of this crisis (UNESCO, 2021; UNICEF, 2021). While a large number of students around the world relied on virtual learning over the internet, for those with no internet access, education came to a full stop (UNICEF et al., 2020). In India specifically, the pandemic resulted in the closure of 1.5 million schools and the nationwide lockdowns in 2020 impacted 247 million children enrolled in elementary and secondary schools (Hindu, 2021). Considering that over six million girls and boys were out of school even before the COVID-19 crisis began, this issue is of particular concern for the country. School closures would likely widen the gap between better-off and worse-off students (Economist, 2021). As such, prioritizing educational recovery is crucial to avoid a generational catastrophe, especially for the poor in rural areas. With this in mind, we examine the causal effects of a low-cost technology based intervention on student learning outcomes during the COVID-19 pandemic in India.

Given that internet access is both limited and inconsistent in India (Gurudu, 2021; Kawoosa, 2020), several forms of distance-learning platforms like television programmes, radio programmes and phone tutorials from teachers were encouraged and utilized by the government to maintain continuity in learning. However, there remains limited evidence of the efficacy of such programs on the learning outcomes of students. In this paper, we explore the impact of a voice-call based intervention that uses text messages and phone calls to impart education to students in primary schools, on their learning outcomes. More specifically, we report results from the evaluation of a Home-Based Learning program launched during the pandemic.

The intervention targeted early-grade learners of government schools in under-resourced communities in three districts in the state of Odisha, India. The program provided remote instructions for teaching basic number recognition and arithmetic operations, and the native language, Odia, via phone call and simple text messages along with automated voice calls to children aged 6–10 years. The program was implemented by ThinkZone, a social enterprise based in Odisha that works towards improving the educational outcomes of children.

Using a difference-in-differences framework, we find that the treatment led to an improvement in learning levels of the student by 4.69 percentage points and 5.52 percentage points for basic number recognition and arithmetic operations (Maths) and Odia respectively. Our results are robust to alternative methods of estimations and matching techniques such as propensity score matching, inverse propensity weighting estimators, and the application of Lee (2009) bounds. The

* Corresponding author.
E-mail addresses: mojha@jgu.edu.in (M. Ojha), kartik@iimnagpur.ac.in (K. Yadav).

https://doi.org/10.1016/j.econedurev.2023.102439
Received 29 June 2022; Received in revised form 2 July 2023; Accepted 5 July 2023
Available online 22 July 2023
0272-7757/

treatment yields important effects for both boys and girls and we note a large increase in the endline examination scores for those students who had scored below the mean in their baseline examination. Our results suggest that well-designed low-cost interventions, particularly for low income countries, are beneficial for students' sustained learning in the face of unexpected shocks.

The remainder of the paper is organized as follows. Section 2 discusses the related literature, followed by Section 3 which provides the background of our study. This is followed by Section 4 which describes the intervention design. In Section 5 we present the estimation strategy and the results of the evaluation in Section 6, followed by heterogeneity analysis in Section 7, robustness in Section 8. Section 9 concludes and discusses the limitations of the analysis.

2. Related literature

Our paper contributes to three strands of literature. First, we add to the broad literature on educational outcomes of children in developing nations based on both experimental and quasi-experimental studies which involve providing learning materials 'at the right level' to students (Afridi, Barooah, & Somanathan, 2020; Bando, Gallego, Gertler, & Fonseca, 2017; Banerjee et al., 2016; Banerjee, Cole, Duflo, & Linden, 2007; Berry, Kannan, Mukherji, & Shotland, 2020; Lakshminarayana et al., 2013; Muralidharan, Singh, & Ganimian, 2019; Muralidharan & Sundararaman, 2011; Naik, Chitre, Bhalla, & Rajan, 2020; Seid, 2016; Yanguas, 2020). For India specifically, extant research finds mixed evidence. While Afridi et al. (2020), Banerjee et al. (2016, 2007) and Lakshminarayana et al. (2013) in general, find positive impacts of providing supplementary remedial teaching and learning materials on learning outcomes, Berry et al. (2020), find of no effect on student achievements.

Second, our paper adds to the research that focuses specifically on technology based educational interventions on student outcomes and students' out-of-school learnings (Aker & Ksoll, 2020; Aker, Ksoll, & Lybbert, 2012; Angrist et al., 2020; Angrist, Bergman and Matsheng, 2020; Azevedo, Hasan, Goldemberg, Iqbal, & Geven, 2020; Bacher-Hicks, Goodman, & Mulhern, 2021; Barrera-Osorio, Gertler, Nakajima, & Patrinos, 2020; Carlana & La Ferrara, 2021; Escueta, Nickow, Oreopoulos, & Quan, 2020; Hossain, 2020; Ksoll, Aker, Miller, Perez, & Smalley, 2015; Muralidharan et al., 2019; Naik et al., 2020; Rodriguez-Segura, 2021; Rodriguez-Segura & Schueler, 2022; Sabates, Carter, & Stern, 2021; Sung, Chang, & Liu, 2016). Largely, the evidence points to improvements in student achievements. For instance, Naik et al. (2020) report increases in the test scores in science, english and mathematics as a result of use of technology in a large-scale experiment in the state of Karnataka, India. Similarly, Sung et al. (2016) quantify the overall effectiveness of integrating mobile technologies into education on student learning achievement. They find larger effects of using mobile devices than using desktop computers or not using mobile devices at all. In consonance with such findings, Muralidharan et al. (2019) present experimental evidence that technology-led instructional program improve test scores in both Hindi and math subjects. In addition, mobile use to enable parents to teach literacy skills to children and monitor their child's progress in school has been documented in Bergman and Chan (2021), Berlinski, Busso, Dinkelman, and Martínez (2021), Doss, Fahle, Loeb, and York (2019), Kraft (2020) and York, Loeb, and Doss (2019).

While considerable evidence exists on the effects of technology use in educational interventions, there remains much to be learnt about the effects of technology based interventions on educational outcomes during the COVID-19 pandemic. Within this recent literature, Angrist, Bergman, Evans et al. (2020), Angrist, Bergman and Matsheng (2020), Crawfurd, Evans, Hares, Sandefur, et al. (2021), Hassan, Islam, Siddique, and Wang (2021), Lichand and Christen (2021), Radhakrishnan et al. (2021) and Schueler and Rodriguez-Segura (2022) specifically evaluate the effects of low-tech solutions that leverage text messages and direct phone calls and home-based/remote learning on students' learning outcomes. This is the third strand of literature to which our paper contributes. We focus on the effects on student outcomes during the pandemic through a technology based Home-Based Learning intervention for India. This is important because India had one of the longest and most stringent lock down in the world (Stringency Index, OxCGRT) with primary schools shut since March 2020. In the absence of any form of internet connectivity, the majority of rural children simply did not study at all, a problem which is compounded by a lack of computers, smart phones and power supply at home. Building on India's dense mobile coverage, our low-cost technology intervention is based on a simple voice-call to impart education in communities that are lacking in other resources.

3. Context

The state of Odisha was hit by COVID-19 cases initially in the month of March 2020, with little respite thereafter. With community transmission setting in, the months from July 2020 to November 2020 saw the rampage of the virus with a rapid spread of infection.[1] As a preventive measure, the state government of Odisha shut all schools and colleges on March 13th 2020 till June 2020 which was further extended over the course of the year till December 2020.[2] In August 2020, the state government announced that it would roll out the Siksha Sampark Yojana (Mohanty, 2020a) to provide e-education with content shared on platforms like television, radio, smartphone amongst others, but there is insufficient evidence of this. According to the ASER (2020) report, only about 10% of children enrolled in classes 1 to 2 and 18% of students in classes 3 to 5 received any learning materials. Overall only 24% of all enrolled children received learning materials from their teachers via WhatsApp and unsurprisingly, private school students received more than twice the learning materials than their government school counterparts.[3] Meanwhile, other initiatives like the E-Vidyalaya App and Madhu App were launched (which mainly focussed on the senior secondary classes) but required a smartphone and an internet connection. However, these initiatives were able to reach only 2.2 million students of the 6 million students targeted during the lockdown. As part of the state government's Odisha Siksha Sanjog Program,[4] students of both private and public schools received learning materials via WhatsApp but almost 50% of students could not access them due to reasons such as lack of smartphones, limited internet access and connectivity issues (ASER, 2020). Therefore, there was a large learning gap for primary school students who attended government schools. Our intervention was targeted at these students with the aim of mitigating their learning loss.

The lockdown lifted in early 2021 for a brief period, however the engagement of students fell further, below 1 million (Barik, 2020). On the directive of the state government in March 2021, students of grades 1 through 8 were promoted without the requirement of any examinations for a second consecutive year (HindustanTimes, 2021). This was followed by a second COVID wave commencing in April 2021, which brought in a second round of school closure and devastation.

[1] Odisha was one of the top ten affected states in the country and clocked a test positivity rate of more than 10 per cent during the month of July (Agragami, 2020).

[2] The Odisha government postponed opening of schools until December 31 despite poor learning outcomes of students as a result of poor network connectivity (Mohanty, 2020b).

[3] The Shiksha Sampark Yojana in Odisha was part of the NITI Aayog project SATH-E, 'Sustainable Action for Transforming Human Capital Education', the first phase of which was completed in March 2020. While phase two of the project in Odisha was announced to be launched in October 2020, it was not operational during the Home-Based Learning intervention period.

[4] For more details on the program, please refer to http://ntse.scertodisha.nic.in/public/sikshya-sanjog-data.aspx.

4. Details of the intervention

4.1. Treatment and control groups

Three districts of Odisha were chosen for the intervention: Cuttack, Bhadrak and Khordha. The intervention districts were chosen such that they represent similar levels of development. None of the three districts share a border with any neighboring state, which ensures that there are no spillovers from public education provisions taking place in the neighboring states. The three districts are similar on indicators such as literacy rates, sex ratio, and work participation rates. In terms of overall literacy rates, Cuttack stands at 85.50%, Bhadrak at 86.88% and Khordha at 82.80%. The female literacy rate was 79.55%, 75.83% and 81.61% for Cuttack, Bhadrak and Khordha, respectively. The sex ratio in Cuttack, Bhadrak and Khordha are also similar at 940, 981 and 929 females per 1000 males, respectively.[5] As per the Census 2011, the work participation rate in Cuttack stood at 35.7%, Bhadrak at 31.1% and 35.2% (Government of Odisha, 2018). The population of Khordha and Cuttack are similar while Bhadrak is slightly smaller. Growth rate of Cuttack was 12.10%, Bhadrak at 12.9% and Khordha slightly higher at 19% as per the Census 2011. The literacy rate among the Scheduled Caste (SC) population is also similar (76.08%, 74.03% and 76.82% for Cuttack, Bhadrak and Khordha, respectively) and higher than the state average of 69.02%. The Gross Enrollment Ratio across the three districts was 97.40% (Cuttack), 98.72% (Bhadrak) and 99.77% (Khordha) according to the OESPA (2019) report. All three districts had a lower than average Scheduled Tribe (ST) enrollment in elementary education, at 6.63% for Cuttack, 4.08% for Bhadrak and 11.10% for Khordha, against the state average of 30%. Female student enrollment was similar at 47.82%, 48.10% and 46.99%, along with similar Pupil-Teacher Ratio (PTR) of 19.99, 20.01 and 20.86 for Cuttack, Bhadrak and Khordha, respectively.[6] However, the state averages for overall literacy rate, and female literacy are at 72.87% and 64.01% respectively, and as such are lower than the educational indicators for the three districts in our sample. The district averages on indicators like socio-economic status, and caste distribution also differ from the state average. Thus our sample cannot be considered as a representation of the entire state of Odisha.

Within each chosen district, some clusters were shortlisted based on the cooperation from the state government during the pandemic with regards to sharing of information such as phone numbers of students enrolled in the government schools, as well as volunteer base of ThinkZone that facilitated the implementation of the intervention on ground. Following this, some clusters in each district were randomly assigned to the treatment group and some to the control group. In the Bhadrak district, the treatment clusters include Karanjamala and Dosingha while the control group comprise the Kuamara cluster and Paikasahi cluster. The treatment clusters in Cuttack are Santapur, Koiladadhibarman and Buhalo Nodal Cluster; and the control clusters are Jharapada, Pearimoni Cluster and Nagaspur. Similarly, for Khordha district, Bajapur and Jagulipatna clusters constitute the treatment group and Jariput and Narangarh clusters, the control group.

Phone numbers for students were obtained from the government schools in each cluster. Out of 6907 phone numbers provided by the schools, about 6369 numbers were found to be valid[7] and these students were onboarded for the intervention. While we acknowledge the concerns related to the validity of phone-based assessments, data collection through in-person surveys were not possible due to severely restricted mobility and social distancing guidelines to contain the spread of COVID-19. With high mobile phone coverage in India as per the NFHS-4, evidence suggests that lack of mobile phone ownership is unlikely a major barrier to collecting data samples (Coffey, Hathi, Khalid, Khurana, & Thorat, 2021). In addition, Ahluwalia et al. (2023), Kumar, Gupta, Ahluwalia, and Gupta (2022) and Rodriguez-Segura and Schueler (2022) provide evidence that establishes the reliability and validity of phone-based assessments and recommend effectively integrating them in teaching–learning pedagogies.[8]

4.2. Intervention design

The objective of the intervention was to improve the educational outcomes of early-grade students in under-resourced communities, using a low-cost technology and remote-instructions model with activity-based methodology and was launched in response to school closures amid the pandemic. The Home-Based Learning program was implemented for all government school students of grades 1 to 5, aged 6–10 years, in the randomly selected clusters. The program provided remote instructions via automated phone calls and simple text messages to engage and teach the children. The voice calls and text messages were sent to a parent's phone number, so that they can help engage their children in the activity-based learning content. Access to the content was provided free of cost to the families.[9]

The intervention was implemented by ThinkZone, a social enterprise that worked with instructors and in consultation with school teachers to deliver the program with support from MoSchool, a Government of Odisha initiative under School and Mass Education Department.[10] The activities were based on the learning outcomes specified by the state government (language and arithmetic skills for primary grades), jointly with the National Council of Educational Research and Training (NCERT).

The details of the intervention are as follows. First, a pre-recorded descriptive audio content was shared with the participants in the treatment group only. The content of the call was based on the learning level of the child, pre-determined from their baseline test score.[11] Second, the same content was also shared via a text messages (*SMS*) along with sensitization messages that highlighted parents' role in children's continuous learning.[12] On the whole, the treatment group received five SMSs and two automated phone calls each week.[13] Third, the automated call and SMS were supplemented by two live phone calls of 15-20 min duration from the instructors each week. On the call, the instructor asked the parent to find the student and put the call on speaker. This arrangement allowed both the parents and children to hear the instructor at the same time and engage in learning. The overall goal was to conduct the activities with both parents and children simultaneously.[14] The instructor confirmed that the child had received

[5] https://www.census2011.co.in/census/state/districtlist/orissa.html.

[6] While we recognize the absence of parallel pre-treatment trends in our analysis, the districts exhibit similar indicators of development to a large extent and are reasonably comparable.

[7] 9.5% of the phone numbers were unreachable, switched off or incorrect in the treatment group and approximately 7.2% in the control group.

[8] Through a small pilot Kumar et al. (2022) document that phone assessments are valid and reliable for measuring students' foundational literacy and numeracy skills.

[9] Learning activities include indoor and outdoor play, puzzles, drawing, drama, and rhymes. We provide examples of the activity-based content in Appendix A.1. of the paper.

[10] https://moschool.in/about-mo-school/.

[11] A baseline test was conducted for both the control and treatment groups in November 2020. A detailed timeline of the intervention is provided in Fig. 1.

[12] A toll-free number was also provided in the voice-based calls and SMS, in case parents/children want to know more about the learning activities. The aim of this communication was to encourage parents to set aside a dedicated time of the day to support children's learning at home.

[13] In sum, this amounted to one participant having received 80 messages, 30 automated calls, and 35–45 live calls from volunteers in the span of four months.

[14] Note that our analysis does not disentangle parental involvement from student performance.

the SMS message sent, and the parent had heard the voice calls. The child was asked questions related to the learning activity.[15]

The live phone calls served to provide additional learning support and motivation to the participants, and increase accountability. This strategy maximized the probability that the child received educational support and lowered future barriers to entry for parents to continue engaging in educational activities. Finally, following ILO (2018) report, for comparability, a similar test to the baseline was repeated for endline examination to capture the causal improvement in student learning level, if any, due to the intervention. The ILO (2018) report notes that it is best to keep follow-up questions and the order of questions as similar to the baseline survey as possible to ensure that they are comparable.[16] It is important to note that while the levels of learning assessed in the endline examinations remained the same as in the baseline test, the questions were not identical.[17] This alleviates concern with regards to whether the intervention taught to the test or coached the students. That said, given that the intervention is concerned with very basic skills in Maths and Odia, teaching to the test is not a severe concern.

4.3. Training of the instructors

Owing to the pandemic, we were severely constrained in carrying out the intervention. To begin with, we were unable to conduct face-to-face training programs of the instructors and hence their identification and training was conducted by ThinkZone. All the instructors were female and belonged to the age-group of 18–35 years with either a graduation degree or were pursuing their graduation. Most instructors had some informal teaching experience. These instructors were randomly assigned to students.

The instructors took a one-week (2 h per day) online pre-program training since they were responsible for making the calls. They were trained to implement various features of the Home-Based Learning program, such as how to interact with parents and children together, how to ask directed questions to children about the learning activities, how to conduct the baseline/endline test, how to use the mobile-based dashboard for monitoring and reporting the performance of children. The training also focused on up-skilling the instructors on how to use technology tools for providing support to children and how to involve children using activity-based teaching methodology. In addition, all instructors took a monthly online training session (2 h/day) focused on how to effectively engage children and parents in the Home-Based Learning program and how to track the monthly performance of children.

The instructors were assigned reporting managers who had teaching experience and were responsible for monitoring the Home-Based Learning program.[18] The reporting managers were trained to use the mobile-based tracking dashboard for monitoring the instructors and to solve the operational challenges faced while implementing the Home-Based Learning program.

4.4. Treatment uptake

The program was limited to government school students. The baseline test was conducted on 3206 students from the treatment group while the endline examination was conducted for 2861 students, with 345 dropouts. Within the control group, the baseline test was taken by 3098 students and the endline by 2386 students respectively. This was due to the fact that several families started to resume economic activity as lockdown restrictions eased and could not be contacted for the endline. The baseline examination was conducted during the first half of November 2020 with the intervention beginning on November 15th, 2020 and ending in the last week of March 2021. Endline examination was subsequently conducted over the phone for both the treatment and control groups during April 2021. Fig. 1 presents a detailed timeline of the intervention.

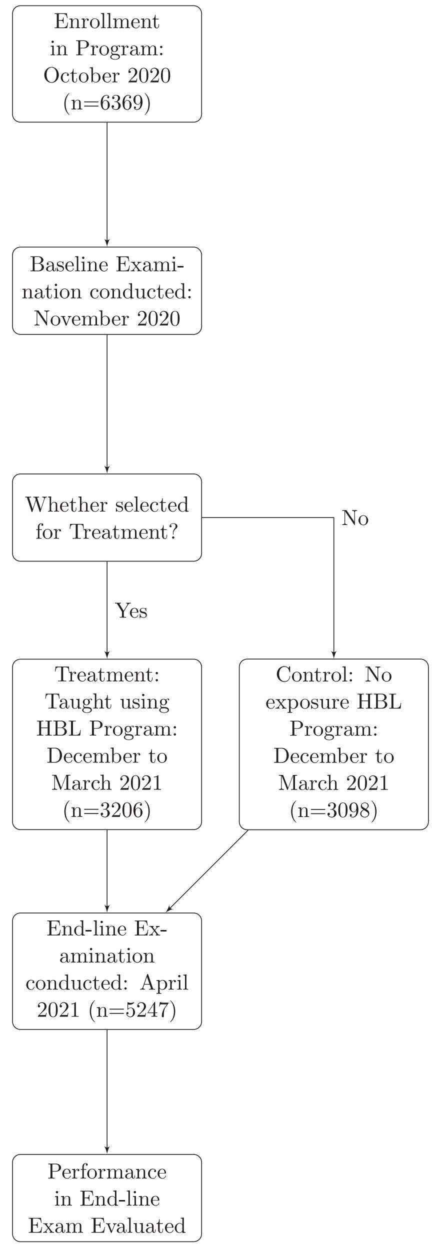


Fig. 1. Timeline for implementation of home based learning program.

4.5. Examination

A baseline test was administered to both the treatment and control groups and learning level scores were assigned to each student based on their performance on the test. The tests were adaptive in nature, and were individually administered to all students over the phone. Both baseline and endline tests assessed students on their Maths and Odia

[15] The four components of the intervention — text messages, automated phone calls, toll-free number calling, and bi-weekly live phone calls were in sync with each other.

[16] Unless there was a major issue with a question in the baseline survey, it is best to leave the wording unchanged in follow-up surveys.

[17] We provide one sample of the Maths test in the Appendix A.3. of the paper.

[18] There was 2-day (3 h/day) training for the reporting managers which covered all the features of the Home-Based Learning program.

Table 1
Sample descriptive statistics.

Variable name	Treatment group Percent/Mean (SD)	Control group Percent/Mean (SD)	[1] *t*-value	[2] *t*-value	[3] *t*-value
DV: Mathematics					
Overall	1.911 (1.049)	1.909 (1.067)	−0.04		
Level 1	45.70%	47.48%			
Level 2	29.54%	26.31%			
Level 3	14.85%	16.07%			
Level 4	7.80%	8.01%			
Level 5	2.12%	2.13%			
DV: Odia					
Overall	2.186 (1.042)	2.138 (1.107)	−1.79		
Level 1	31.25%	35.28%			
Level 2	33.03%	32.05%			
Level 3	22.74%	19.92%			
Level 4	11.73%	9.07%			
Level 5	1.25%	3.68%			
IDV(s)					
Females	0.533 (0.498)	0.522 (0.499)	−0.88	−0.15	−2.11
Age	8.09 (1.399)	8.07 (1.394)	−0.71	−14.35	−8.18
Hindus	0.906 (0.291)	0.905 (0.293)	−0.13	1.45	1.28
Reserved category	0.224 (0.417)	0.196 (0.397)	−2.63	0.63	−3.33
Household size	5.83 (1.905)	5.74 (1.544)	−2.08	−0.13	−0.14
Edu- HH head	5.551 (2.215)	5.525 (1.584)	−0.51	−0.20	−2.97
Own smartphone	0.847 (0.359)	0.859 (0.347)	1.32	−0.34	3.33
Occ-Agri	0.533 (0.498)	0.596 (0.490)	5.11	−1.47	2.43

Note: [1] measures the difference between Treatment and Control group at baseline level, [2] measures the difference for the Treatment group between baseline and endline, [3] measures the difference for the Control group between baseline and endline. Reserved Castes refer to any social category eligible for reservations in India, that is: Scheduled Tribes, Scheduled Castes and Other Backward Classes. Occ-Agri refers to the occupation of the household primarily being agriculture.

language learning levels. The test was the same for all students irrespective of their age, since it covered the spectrum of teaching activities specified by the state government of Odisha with themes suggested by the National Council of Educational Research and Training (NCERT) national alternative learning calendar. The questions were suggested and vetted by the State Council of Educational Research and Training, Government of Odisha and incorporated features of Item Response Theory (IRT) that is typically used to calibrate, and evaluate questions for assessing student ability non-linearly.

The tests broadly covered themes that are aligned to class 1–5 level state textbooks. These tests were designed based on the ASER testing tools that are simple, quick to administer and considered largely reliable in the context of India (Vagh, 2012).[19] According to the ASER (2018) report, only about 50 percent of students enrolled in grade 5 could read a grade 2 level text. With concerns about low learning levels even for older children, the examinations were designed regardless of age and instead took into consideration the levels of mastery of students on different concepts. The test focussed on understanding what the students can do and the skills they have mastered. Accordingly, the test classified the students at different levels based on their performance.[20] More specifically, Maths and Odia skills were tested on *five* levels. Within the question paper itself, questions were asked in increasing order of complexity and the levels were not mutually exclusive.

For Maths, the first level was based on numbers, level 2 on addition, level 3 on subtraction, level 4 on multiplication and level 5 on division. On similar lines, for Odia, level 1 was *abana* (basic alphabet test), level 2 was *matra* (advanced phonics), level 3 was *phala* (complex words), level 4 was *Yuktakhara* (words with conjunction), and level 5 was reading comprehension. The test was designed in the following manner. Each level consisted of 4 questions, and to pass one level, a student has to answer all four questions. If the student answers three questions successfully, they will remain at the same level and if two questions are answered, then the student will be below that level. All students are asked questions starting from level 1 (the easiest level) followed by questions from the next level and so on. As a specific hypothetical example: consider Rashida in Class 4 answers 4 questions in level 1 of Maths. She then moves to level 2 questions and answers 4 questions. Then she moves to level 3 but only answers 2 questions. Rashida is then assigned level 2 for Maths.

5. Data and methodology

5.1. Data

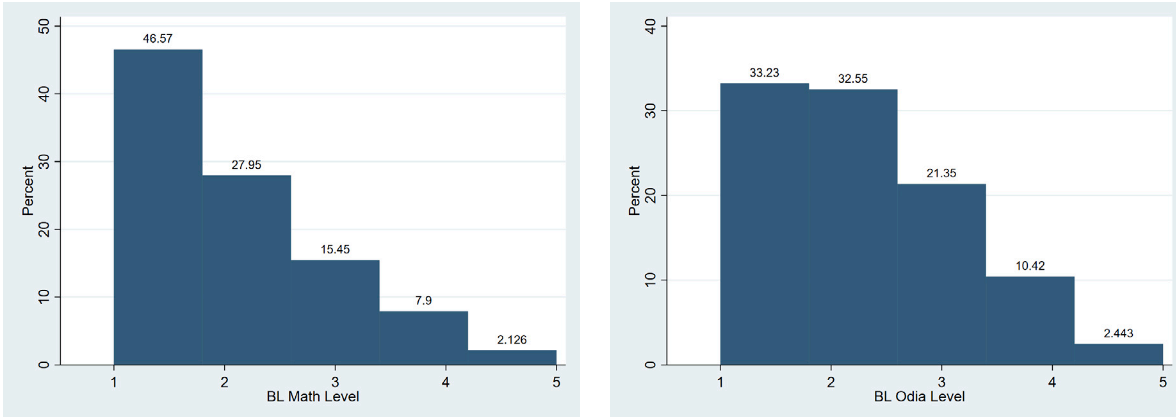
Our outcome variables are endline examination scores in $Maths$ and $Odia$ respectively.[21] The first set of scores take on a range of values from 1 to 5 denoting the different *levels* of learning for students. Next, we also measure our scores as *percentage correct answers* and finally, we use a binary indicator of *improvement*, which takes a value '1' if the student moves from a lower level in the baseline to a higher level in the endline examination, and '0' otherwise.

Our main explanatory variable is the indicator for the treatment, $Treatment$, that takes value 1 if the student is assigned to the Home-Based Learning program and 0 otherwise. Our set of control variables at the child level include age and gender. The controls at the household level are religion (Hindu or non-Hindu), caste (Scheduled Caste, Schedule Tribes, Other Backward Caste or General), the indicator for number of years the household has lived in the same location, size of the household, educational attainment of the father and mother, indicators for whether the household owns assets such as TV, radio, smartphone, agricultural land, house, livestock, and indicator for occupation of the household (agriculture or non agriculture). Table 1 presents the summary statistics of our variables, while also providing a comparison of the treatment and the control sample.

[19] These tests focus on child's level of foundational reading skills (letter identification, word decoding, and so on) and basic math ability (number recognition, subtraction, and division) (Vagh, 2012).

[20] We provide a sample test in Appendix A.3. of the paper. More details on design of such tests can be found at http://www.asercentre.org/, which heavily guided the tests conducted.

[21] We refer to basic number recognition and arithmetic operations as Maths in our model, results and tables.



(a) Distribution of baseline Maths Score (levels) (b) Distribution of baseline Odia Score (levels)

Fig. 2. Student baseline scores in levels.

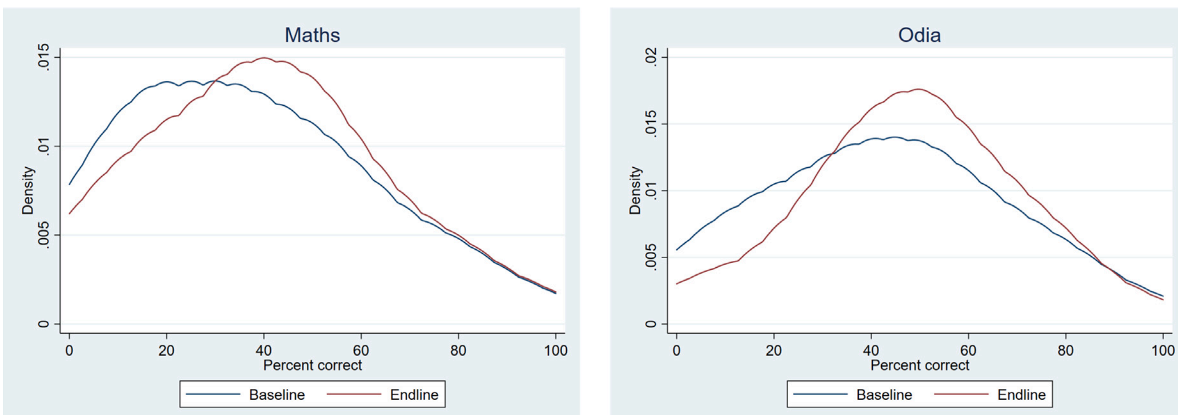


Fig. 3. Distribution of percentage correct scores.

Figs. 2(a) and 2(b) present the distribution of scores in levels for the combined sample in $Maths$ and $Odia$ before the intervention began. We also present the distribution of percentage of correct responses in $Maths$ and $Odia$ for the baseline and endline examination in Fig. 3. As can be seen, the distribution has shifted rightwards for both $Maths$ and $Odia$ from baseline to endline and the average achievement measured shows an improvement. Next, we present the scatter-plots to illustrate the changes in the scores in levels for control and treatment groups, following the intervention, in Figs. 4 and 5. Here, we see that the students in the treatment group show a greater improvement than their counterparts in the control group.[22] We further present sub-plots of percentage correct answers, by round and treatment status in Fig. 6. We find that while the baseline scores are comparable, the increase in the endline scores is driven largely by the higher average score of the treatment group (see Fig. 6).

5.2. Estimation strategy

We begin by using the following estimation strategy to study the treatment effect of the Home-Based Learning intervention on the scores for $Maths$ and $Odia$:

$$Y_{ict} = \beta_0 + \beta_1 Treatment_c + \beta_2 Post_t + \beta_3 Treat_c * Post_t + \beta_4 \mathbb{X} + \lambda_d + \epsilon_{ict} \quad (1)$$

where Y_{ict} refers to the $Maths$ and $Odia$ scores of student i in cluster c at time t with values ranging from 1 to 5[23]; $Post_t$ takes value 0 for periods before the intervention and 1 for periods after the intervention; $Treatment_c$ is a dummy variable that takes value 1 for the treated cluster and 0 for those not treated. $Treat_c * Post_t$ is a binary variable equal to 1 if the observation is from the endline and is for a treated unit. $\mathbb{X}$ are the vector of individual and household level controls; λ_d denotes the district fixed effects and ϵ_{ict} summarizes the influence of all other unobserved variables. Our parameter of interest, β_3, essentially captures the change in students' endline examination scores due to the treatment. Inclusion of district fixed effects addresses the concern of any time invariant district level characteristics (such as administrative efficiency) that may be correlated with the treatment in our estimation. Given that we have few clusters in our intervention, we utilize Roodman, Nielsen, MacKinnon, and Webb (2019)'s wild cluster bootstrapping procedure for inference purposes to alleviate any concern regarding the small number of clusters in our intervention.[24] The wild cluster bootstrap is especially useful when conventional inference methods are unreliable because large-sample assumptions do not hold. By default, it generates

[22] The arrows show the direction of change of scores from baseline to endline.

[23] We acknowledge the condition that estimating the average treatment effect on the treated requires the parallel trends assumption, which in turn imposes linearity on the outcome variable, that is the test scores. Thus, we consider the percentage correct answers as the outcome.

[24] This is different from the usual bootstrap method which may not perform well in difference-in-differences cases where the number of treated clusters is too few.

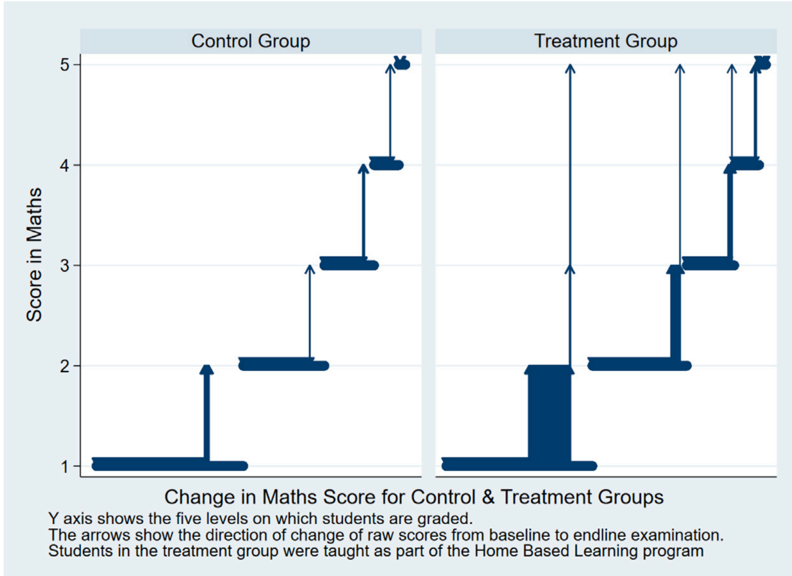


Fig. 4. Improvement in Maths Score (levels) for students.

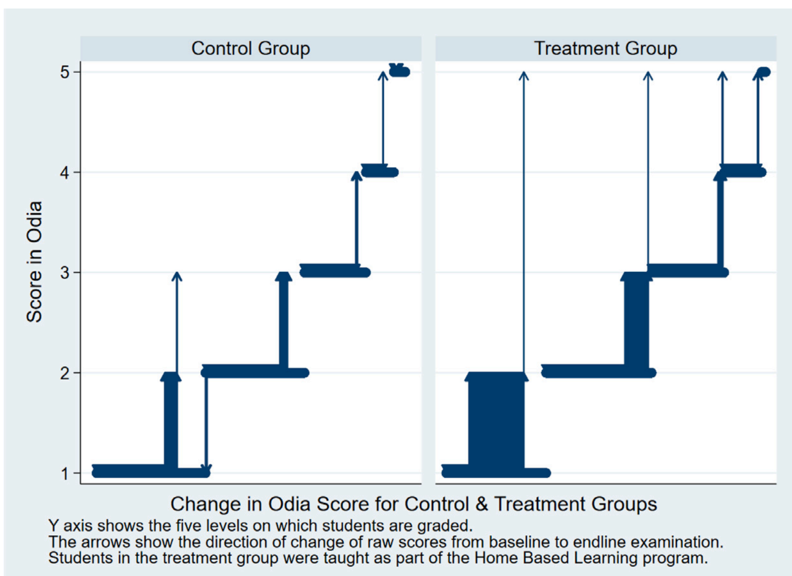


Fig. 5. Improvement in Odia Score (levels) for students.

999 wild cluster bootstrap samples using the Rademacher distribution with the null hypothesis imposed and provides a way to improve inference with few number of clusters. The algorithm does not produce standard errors and instead, the inference is based on p-values and confidence sets (Roodman et al., 2019). As such, we present the wild bootstrap confidence intervals in parenthesis below the estimates. We also present the *p*-value from the Wald test at 95% level of confidence.

Further, we extend our analysis and estimate a model where the scores are measured as percentage of correct answers obtained by students in the examination. As before, we employ the wild cluster bootstrap procedure for inference (Roodman et al., 2019).

Finally, we estimate a model where we convert the outcome into a binary variable, *improvement*, which takes the value 1 if there is any improvement shown by the student in the endline test, as compared to her baseline test score in levels. Therefore, *improvement* equals 1 if endline score - baseline score ≥1 and 0 otherwise. We first estimate Eq. (2) using a Linear Probability Model (LPM), followed by a Probit estimation model:

$$\triangle Y_{ic} = \alpha_0 + \alpha_1 Treatment_c + \alpha_2 \mathbb{X} + \eta_d + \epsilon_{ic} \tag{2}$$

Improvement from level 1 to 2 is regarded the same as an improvement from level 2 to 3, and so on. Our assumption that categories are equally spaced does not dilute our goal, i.e. to estimate the average shift in the test scores. As before, we present inference using Roodman et al. (2019)'s wild cluster bootstrap.

6. Results

6.1. Scores in levels

Table 2 presents the results for the effect of the Home-Based Learning intervention on the test scores of $Maths$ and $Odia$, considered as levels in the range of 1 to 5. For both $Maths$ and $Odia$, column (1) is a naive specification with no controls. We subsequently add the full set

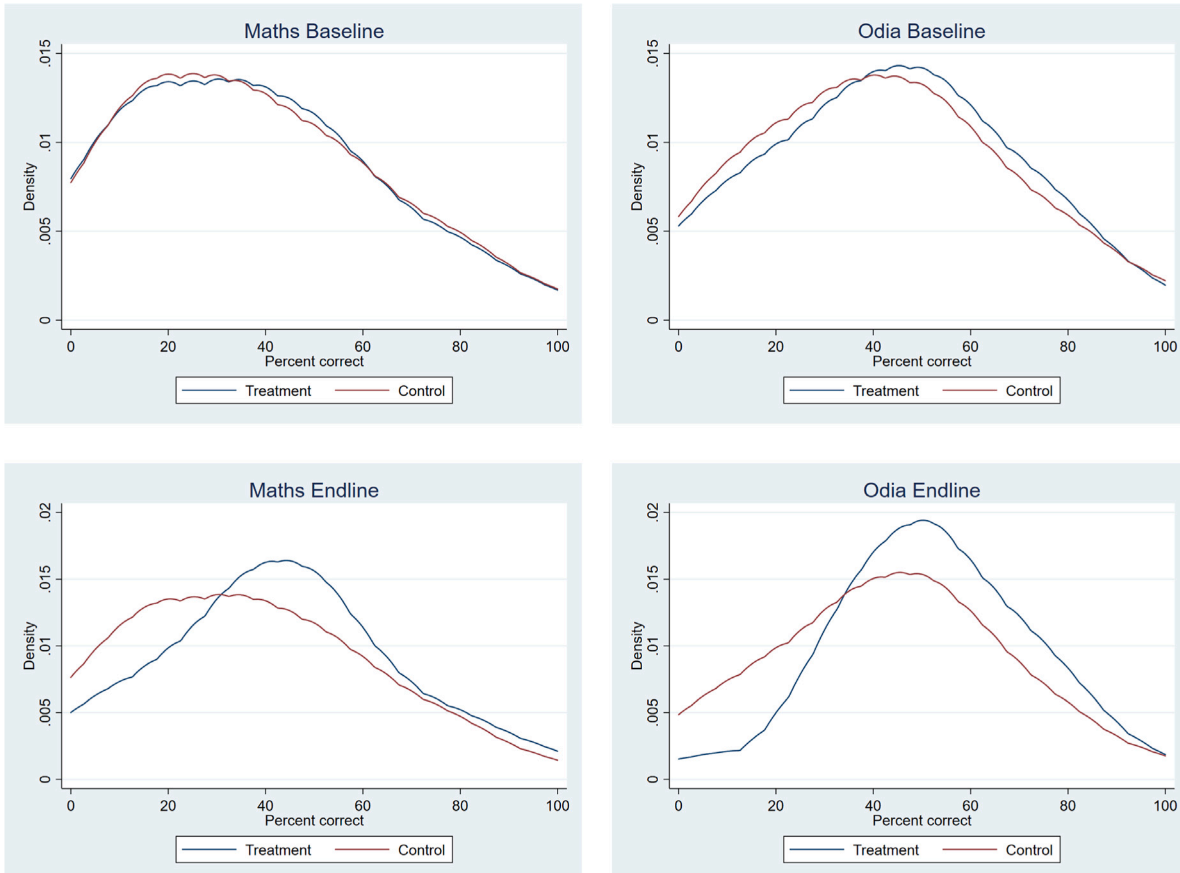


Fig. 6. Distribution of percentage correct scores, by round and treatment status.

Table 2
OLS estimation for levels: Difference-in-differences results for Maths and Odia.

	Mathematics			Odia		
	(1)	(2)	(3)	(1)	(2)	(3)
Treat × Post	0.1736**	0.1744**	0.1744**	0.206***	0.226***	0.226***
Wild cluster bootstrap *p*-value	0.012	0.011	0.014	0.001	0.001	0.002
Wild cluster bootstrap CI	[0.0419, 0.3013]	[0.0397, 0.2975]	[0.0409, 0.304]	[0.1055, 0.2802]	[0.141, 0.2995]	[0.1277, 0.2933]
Observations	10,494	9584	9584	10,494	9584	9584
R-squared	0.012	0.095	0.097	0.028	0.214	0.218
Demographic characteristics	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes

Note: Outcome variable in all specifications is the endline score in Maths and Odia that ranges from 1 to 5. Col. (1) presents the OLS estimation results without controls. In columns (2) and (3), we sequentially add individual controls and the district fixed effects. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap *p*-value and confidence intervals are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

Table 3
OLS estimation for percent score: Difference-in-differences results for Maths and Odia.

	Mathematics			Odia		
	(1)	(2)	(3)	(1)	(2)	(3)
Treat × Post	4.677**	4.700***	4.693**	5.291***	5.530***	5.524***
Wild cluster bootstrap *p*-value	0.012	0.009	0.018	0.000	0.001	0.001
Wild cluster bootstrap CI	[1.511, 7.691]	[1.398, 7.765]	[1.248, 7.876]	[3.012, 7.141]	[3.285, 7.716]	[3.406, 7.852]
Observations	10,494	9584	9584	10,494	9584	9584
R-squared	0.012	0.091	0.092	0.030	0.199	0.203
Demographic characteristics	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes

Note: Outcome variable in all specifications is the percentage score in Maths and Odia that ranges from 0 to 100. Col. (1) presents the OLS estimation results without controls. In columns (2) and (3), we sequentially add individual controls and the district fixed effects. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap *p*-value and confidence intervals are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

Table 4
Improvement: Linear probability model estimation results for Maths and Odia.

	Mathematics			Odia		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	0.168**	0.169***	0.178***	0.200***	0.229***	0.226***
Wild cluster bootstrap p-value	0.012	0.003	0.001	0.003	0.003	0.000
Wild cluster bootstrap CI	[0.0462, 0.2919]	[0.0494, 0.2812]	[0.0504, 0.2909]	[0.0783, 0.2816]	[0.1129, 0.3056]	[0.0946, 0.314]
Observations	5247	4792	4792	5247	4792	4792
R-squared	0.066	0.078	0.093	0.061	0.116	0.117
Demographic characteristics	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes

Note: Outcome variable in all specifications is the binary variable 'Improvement' which takes a value, 1 if the endline level score of the student is higher than or equal to the baseline level score in Maths and Odia, respectively and 0 otherwise. Col. (1) presents the LPM estimation results without controls. In columns (2) and (3), we sequentially add individual controls and the district fixed effects. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap p-value and confidence intervals are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

Table 5
Improvement: Probit estimation results for Maths and Odia.

	Mathematics			Odia		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	1.058**	1.102***	1.153***	0.776***	0.976***	0.962***
Wild cluster bootstrap p-value	0.010	0.002	0.000	0.001	0.000	0.001
Marginal effect	0.190***	0.202***	0.207***	0.204***	0.241***	0.237***
	(0.056)	(0.055)	(0.047)	(0.040)	(0.039)	(0.039)
Observations	5247	4792	4792	5247	4792	4792
Demographic characteristics	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes

Note: Outcome variable in all specifications is the binary variable 'Improvement' which takes a value, 1 if the endline level score of the student is higher than or equal to the baseline score in Maths and Odia, respectively and 0 otherwise. Col. (1) presents the probit estimation results for without any controls. In col. (2) and (3), we sequentially add individual controls and the district fixed effects. In all columns, we report the marginal effects of the treatment on our outcome. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap p-value are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

of controls and district fixed effects in columns (2) and (3) respectively. Column (3) is our most preferred specification.

Overall, our results in Table 3 show that being a part of the Home-Based Learning program increases the test scores in the endline examination for both $Maths$ and $Odia$ compared to the cohort which was not a part of the program. We start by looking at the results from the naive specification in column (1) for both $Maths$ and $Odia$. In the absence of any controls, the average $Maths$ level of students increase by 0.173 units, and $Odia$ by approximately by 0.206 units. The wild cluster bootstrap p-values associated with these specifications show that the effects are statistically significant. The effects remain statistically significant upon inclusion of demographic controls. Finally, from our most preferred specifications, including all the demographic controls as well as district fixed effects in column (3), we note a treatment effect of 0.174 units for $Maths$ level scores and 0.226 units for $Odia$ level scores for the treated cohort compared to those who were not a part of the program. The wild cluster bootstrap p-value and the test statistic in our preferred specifications also indicate that our results are statistically significant.

6.2. Percent scores

Table 3 presents our results from the OLS estimation for $Maths$ and $Odia$ of Eq. (1) where the scores represent the percentage correct answers obtained by the students on the examination. As before, for both $Maths$ and $Odia$, column (1) includes no controls, column (2) includes the full set of controls and column (3) presents the preferred specification with additional district fixed effects.

In line with our results above, we find that the Home-Based Learning program leads to an increase in the $Maths$ and $Odia$ scores. From our preferred specification in column (3), with all controls and district fixed effects, we note an increase of 4.69 percentage points (pp) in $Maths$ scores and 5.52 pp in $Odia$ scores. The wild cluster bootstrap confidence intervals and p-value show that the effects are statistically significant.

6.3. Improvement

Table 4 presents the results from the LPM estimation of Eq. (2). We add our controls in the same sequence for both $Maths$ and $Odia$ as in Table 3. Column (3), including the full set of controls and district fixed effects, is our preferred specification for both the subjects. We find that the treatment leads to 17.8 pp increase in probability of improvement in the endline $Maths$ score for students in the treated cohort and a 22.6 pp increase in $Odia$ scores respectively. The wild cluster bootstrap p-values show that these effects are significant.

Next, Table 5 reports the marginal effects of the treatment from a Probit estimation of Eq. (2). The results are in the same progression as in Table 3. The marginal effects in column (3) are comparable to the estimates from the LPM results in Table 4. For $Maths$, the marginal effects shows a 20.7 pp rise in the probability of improvement for the treated cohort. Similarly for $Odia$, we find that the treatment leads to a 23.7 pp increase in the likelihood improvement for the treated cohort relative to the control group. The results are statistically significant based on the wild cluster bootstrap p-values.

7. Heterogeneity analysis

To the extent that average effects may mask interesting heterogeneous effects of the treatment, we cut our analytical sample in different ways and examine whether the treatment effects vary across these subsamples. We utilize OLS estimation techniques with all controls and district fixed effects to examine the treatment effect on the scores in levels, and as percentage correct answers, and Probit estimation for the improvement outcome respectively. Our subsamples are based on: *gender* and *student's baseline scores*.

By gender: Table 6 presents the results for $Maths$ and $Odia$ in Panel A and B respectively. Column (1) presents the OLS results for scores in levels followed by scores as percentage correct answers in column (2), and the Probit marginal effects in column (3). At the outset, we note that the treatment has a statistically significant effect on both girls and boys in the treated cohort. The effects are larger in magnitude for

Table 6
Heterogeneity: Treatment effect by Gender.

	Panel A: Maths					
	Female students			Male students		
	[1]	[2]	[3]	[1]	[2]	[3]
Treat × Post	0.191***	5.215***	1.447***	0.157**	4.121**	0.942**
Wild cluster bootstrap p-value	0.002	0.000	0.000	0.044	0.017	0.017
Wild cluster bootstrap CI	[0.064, 0.3012]	[0.14, 0.3873]		[0.0021, 0.2881]	[0.0919, 0.337]	
Marginal effect			0.251***			0.170***
			(0.044)			(0.052)
Observations	5018	5018	2509	4566	4566	2283
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B: Odia					
	Female students			Male students		
	[1]	[2]	[3]	[1]	[2]	[3]
Treat × Post	0.234***	5.506***	0.968***	0.217***	0.270***	0.977***
Wild cluster bootstrap p-value	0.000	0.000	0.001	0.002	0.001	0.001
Wild cluster bootstrap CI	[0.1551, 0.2948]	[0.2577, 0.3754]		[0.1102, 0.2971]	[0.1933, 0.3399]	
Marginal effect			0.245***			0.231***
			(0.038)			(0.041)
Observations	5018	5018	2509	4566	4566	2283
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table presents the results by gender for both Maths and Odia. In column [1], the outcome is measured in terms of level, ranging from 1 (lowest) to 5 (highest). In columns [2], we show the percentage correct answers that range from 0 to 100. Columns [1] & [2] are estimated using OLS. In column [3], the outcome is a binary indicator which takes value 1 if the student shows any improvement in end line score in Maths and Odia, and 0 otherwise. The specifications in column [3] are estimated using probit estimation and marginal effects are reported. All specifications include full set of controls and district fixed effects. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap p-value and confidence intervals are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

Table 7
Heterogeneity: Treatment effect by students' baseline test score.

	Panel A: Maths					
	Below mean			Above mean		
	[1]	[2]	[3]	[1]	[2]	[3]
Treat × Post	0.218**	5.963***	1.363***	0.053	1.606	0.665**
Wild cluster bootstrap p-value	0.015	0.009	0.000	0.242	0.375	0.017
Wild cluster bootstrap CI	[0.0497, 0.3659]	[1.964, 10.22]		[−0.03643, 0.116]	[−1.464, 3.293]	
Marginal effect			0.263***			0.067***
			(0.053)			(0.038)
Observations	7182	7146	3591	2402	2438	1201
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B: Odia					
	Below mean			Above mean		
	[1]	[2]	[3]	[1]	[2]	[3]
Treat × Post	0.341***	8.532***	1.156***	0.045**	1.009**	0.864***
Wild cluster bootstrap p-value	0.001	0.001	0.001	0.016	0.044	0.001
Wild cluster bootstrap CI	[0.2322, 0.4219]	[5.464, 11.27]		[0.0075, 0.0705]	[0.0695, 1.706]	
Marginal effect			0.335***			0.068***
			(0.041)			(0.020)
Observations	6248	6213	3124	3336	3371	1668
Demographic characteristics	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table presents the results by students' scores (below and above mean) in the baseline test for both Maths and Odia. In column [1], the outcome is measured in levels ranging from 1 (lowest) to 5 (highest). Column [2] reports the results using the percentage correct answers that range from 0 to 100. Columns [1] & [2] are estimated using OLS. In column [3], the outcome is a binary indicator which takes value 1 if the student shows any improvement in end line score in Maths and Odia endline test and 0 otherwise. The specifications in column [3] are estimated using probit estimation and marginal effects are reported. All specifications include full set of controls and district fixed effects. Inference is based on Roodman et al. (2019)'s wild cluster bootstrapped procedure. Wild cluster bootstrap p-value and confidence intervals are provided. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

girls relative to boys although we do not conclude that the effects are statistically different from each other.

Specifically, for $Maths$, the OLS estimates show that the treatment results in 0.19 units increase in levels for girls and 0.15 units increase in levels for boys. In terms of percentage correct answers, the treatment leads to a statistically significant rise by approximately 5.2 pp for girls and 4.1 pp for boys. We also find statistically significant marginal effects from the probit estimation for both subsamples. The marginal effects indicate a 25.1 (17) pp increase in the likelihood of improvement in $Maths$ scores for girls (boys) in the treated cohort.

On similar lines, we document statistically significant effects for girls and boys in the treated cohort for $Odia$ from the OLS estimations. In terms of levels, we note a 0.23 units increase in the score for girls and 0.21 units increase for boys' score in the treated cohort. The treatment leads to 5.5 pp increase in the percentage correct answers for girls and 0.27 pp increase for boys. Probit marginal effects tell a similar story,

Table 8
Probit estimation: Correlates with attrition.

Variables	Full sample	Treatment group	Control group
Treatment	−0.291***		
	(0.070)		
Baseline Maths	0.007	−0.022***	
	(0.007)	(0.007)	
Baseline Maths × Treatment	−0.038***		
	(0.121)		
Baseline Odia	0.000	−0.023***	
	(0.008)	(0.008)	
Baseline Odia × Treatment	−0.030***		
	(0.134)		
Age	0.022***	0.047***	0.030***
	(0.004)	(0.004)	(0.007)
Age × Treatment	0.041***		
	(0.007)		
Observations	6290	3206	3084
Demographic characteristics	Yes	Yes	Yes
District FE	Yes	Yes	Yes

Note: The table presents the results from a regression of attrition on the covariates. The dependent variable is an indicator for attrition capturing students who were enrolled in baseline (and thus have a baseline score) but were not present for the end-line examination. The specifications are estimated using probit estimation and marginal effects are reported for the full sample with treatment interaction, treatment group and control group respectively. Only those covariates that are statistically significant are presented. Clustered standard errors at the cluster level are presented in parentheses. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

with a 24.6 (23.1) pp rise in the probability of improvement in $Odia$ scores for girls (boys) in the treated cohort. Across the estimations, we document that these effects are statistically significant.

By student ability: Table 7 presents the results capturing whether the effect of the treatment is different for those who scored below and above the mean in the baseline examination for $Maths$ and $Odia$ in Panel A and B respectively. As before, we first present the OLS results for scores in levels followed by the OLS coefficients for scores measured as percentage correct answers and lastly the marginal effects from Probit estimation. Both OLS and probit estimations show a large and statistically significant treatment effect for students who scored below mean in their baseline tests indicating the effectiveness of the treatment.

In terms of levels and percentage, we notice a statistically significant increase for only one sub-sample of students. The treatment leads to a 0.21 units increase in $Maths$ level scores for students below the mean in baseline and a 5.9 pp increase in the $Maths$ percentage correct scores for this sub-sample. The effects are not statistically significant for the sample of students with scores above mean in the baseline examination. The Probit marginal effects show that the treatment leads to a 26.3 pp increase in the $Maths$ score for the below mean students and 6.7 pp increase for those above mean.

For $Odia$, we see a large effect of the treatment for students who scored below the mean in the baseline examination. For scores in levels, we note an increase by 0.34 (0.04) units for below mean (above mean) students. For scores in terms of percentage correct answers, the increase is of the order of 8.5 (1) pp for the below mean (above mean) students and the effect is statistically significant. The marginal effect from probit estimation shows that the treatment leads to 33.5 (6.8) pp increase in the likelihood of improvement in endline scores for below mean (above mean) students.

8. Robustness

To address concerns about non-random selection of sample in our analysis and given the imbalance in the covariates between attriters and non-attriters, we conduct a set of robustness checks.

8.1. Attrition

First, we analyze the determinants of attrition. Table 8 shows results from three probit regressions where the dependent variable is a dummy for attrition. We include all the covariates viz. the treatment status, student's baseline scores (for Maths and Odia), gender, age, parental education and household characteristics. Column (1) presents the results for the entire sample.[25] Our objective through this exercise is to identify the determinants of attrition in the intervention, and whether these differ significantly for the treatment and control groups. We do this by interacting all our observable characteristics with the treatment.

We find that being a part of the treatment group significantly reduces dropouts. This is perhaps driven by the control group having little to gain from answering the endline questions during the pandemic while the treatment group was regularly being contacted and followed-up with. The differential attrition in our students is 12.2 pp with 22.98 pp for control group versus 10.76 pp for the treatment group. We also find that those with higher baseline scores have a lower probability to attrit, after controlling for all other factors. Lastly, we find that older students have a higher probability to drop out. All other observables such as gender, parental education and household characteristics (and their interactions with the treatment) are insignificant in explaining attrition.

Since the attrition between the treatment and control group are not of similar magnitudes, we also compare the attriters' observable characteristics for treatment and control sample separately in columns (2) and (3). The only significant determinant is the age of the student for both the treatment and control groups. The marginal effect is positive indicating that older children are more likely to attrit from both treatment and control groups. This is perhaps driven by the fact that the older students are already well-versed in the concepts taught and thus the program did not sustain their interest.

[25] We only present results for the statistically significant covariates in the table.

Table 9
Lee bounds estimates for the home-based learning program on student learning outcomes.

	Mathematics			Odia	
	Score		Improvement	Score	Improvement
	[UT]	[T]	[UT]	[UT]	[UT]
Lower	−0.048	0.096***	0.090***	0.040**	0.087***
	(0.031)	(0.010)	(0.030)	(0.014)	(0.013)
Upper	0.435***	0.209***	0.508***	0.198***	0.246***
	(0.032)	(0.010)	(0.027)	(0.009)	(0.011)
Trimming proportion	0.137	0.137	0.137	0.137	0.137
Effect 95% CI	[−0.100, 0.489]	[0.079, 0.226]	[0.041, 0.553]	[0.016, 0.214]	[0.064, 0.265]

Note: [UT] refers to Untightened Bounds, [T] refers to Tightened Bounds. For Score, the outcome is student performance ranges from 1 to 5 and the specifications are estimated using OLS. For Improvement, the outcome is a binary indicator which takes value 1 if the student shows any improvement in end line score for the subject and 0 otherwise. Bootstrapped standard errors are presented in parentheses and are calculated from bootstrapping with 250 replications. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

Table 10
Propensity score weighted estimates of the home-based learning program on student learning outcomes.

	Kernel [I]		NN (k = 1) [II]		NN (k = 3) [III]		NN (k = 5) [IV]		IPW [V]	
	Maths	Odia	Maths	Odia	Maths	Odia	Maths	Odia	Maths	Odia
Score	0.284***	0.356***	0.248***	0.363***	0.298***	0.357***	0.291***	0.357***	0.275***	0.355***
	(0.028)	(0.032)	(0.033)	(0.052)	(0.038)	(0.043)	(0.035)	(0.038)	(0.026)	(0.029)
Improvement	0.173***	0.224***	0.173***	0.228***	0.175***	0.219***	0.178***	0.224***	0.171***	0.224***
	(0.007)	(0.010)	(0.008)	(0.015)	(0.009)	(0.013)	(0.009)	(0.013)	(0.008)	(0.010)

Note: For Score, the outcome is student performance on an ordinal scale of 1 to 5 and the specifications are estimated using OLS. For Improvement, the outcome is a binary indicator which takes value 1 if the student shows any improvement in end line score for the subject and 0 otherwise. The propensity scores were calculated using the full set of controls and district fixed effects. Bootstrapped standard errors are presented in parentheses and are calculated from bootstrapping with 100 replications. ***, **, and * represent significance at 1 percent, 5 percent and 10 percent, respectively.

8.2. Lee bounds (2009)

Given that the attrition rate in our sample for the control group is 22.98 pp and the treatment group is 10.76 pp, we utilize the methodology proposed by Lee (2009) which presents a bounds estimator that estimates an interval for the true value of the treatment effect in the presence of nonrandom sample selection (Tauchmann, 2014). The bounds estimator trims either the treated or the non-treated observations so that the share of observations with observed outcome is equal for both groups. The key identifying assumption required for using Lee bounds is the monotonicity assumption, that is, the treatment assignment affects sample selection in only one direction. This means that we assume that some students attrited since they were not assigned to the treatment, but no students attrited as a result of being assigned to treatment. This seems plausible in our context.

We present the Lee bounds estimator results in Table 9 for our outcomes in both continuous forms (in levels) denoted by $Score$ and binary denoted by $Improvement$. The columns specified by [UT] are bounds which do not involve any tightening. We begin with examining the bounds for $Maths$. Taking bootstrapped standard errors into account, we find that the lower bound does not significantly differ from zero in column [I]. Thus, we tighten the bounds using covariates that we find have explanatory power for attrition. Tightening yields much narrower bounds for the treatment effect, and the result is presented in column [II]. The remaining bounds in column [III]-[V] are left untightened since the confidence interval for them does not include the value zero. All results are presented with bootstrapped standard errors with 250 replications. Overall, the results from Table 9 are consistent with our main results and suggest that the treatment effect on those that never suffer from attrition is positive and significant.

8.3. Propensity score matching

Next, we utilize propensity score matching (PSM) estimation which relies on the assumption of 'selection on observables', such that, conditional on observable characteristics, students in the treatment group and control group do not systematically differ along unobservable dimensions. While propensity scores are widely used in the matching literature, no single method can be considered to be the best (Frölich, 2004). Thus, we employ two commonly-used matching methods that are kernel density and k-nearest neighbor. The kernel density estimator compares each student in the treatment group to a weighted average of all control group observations, with the weight for each observation in the comparison group inversely proportional to the difference between that observation's estimated propensity score and the propensity score of the treatment student. In the k-nearest neighbor approach, each student in the treatment group is matched with k students in control group who have the most similar propensity scores. Here, we use three alternative values of k, that are 1, 3, and 5 (Silverman, 1986).

We present the results in Table 10. Panel [I] presents the PSM estimation results using kernel method and the Panels [II], [III], and [IV] present the k-nearest neighbor method for our outcomes in both continuous (*Score*) and binary (*Improvement*) forms. For both matching methods, we use all our controls to compute the propensity scores. The kernel matching results are in line with our OLS results and point to a statistically significant effect of the Home-Based Learning program on the test scores for $Maths$ and $Odia$. Similarly we find statistically significant effect of the treatment on the improvement indicators of the student for both $Maths$ and $Odia$. The k-nearest neighbor matching estimates in Panels [II]-[IV] are largely similar to the kernel matching estimates. Finally, like the OLS and the LPM/Probit results, we find that the treatment effects are stronger for $Odia$.

8.4. Inverse propensity weighting estimator

We also utilize the inverse propensity weighting (IPW) estimator proposed by Hirano, Imbens, and Ridder (2003), which is an alternative to matching estimators, but relies on estimating the propensity score. In calculating the treatment effects, the estimator weights observations by the inverse of nonparametric estimates of the propensity score, rather than the true propensity score.[26]

We present the IPW estimates in Panel [V] of Table 10. These results are also consistent with the OLS results and the Probit results in our main analysis for the continuous $Score$ outcome and the binary $Improvement$ outcome respectively. The results are also largely similar to the PSM results in Panels [I]-[IV] of Table 10.

8.5. Falsification

A possible concern with our main estimation procedure is that the effect picked up by our regressions is not necessarily that of the Home-Based Learning program. Based on our identification strategy, we expect that such a result would not be obtained if we randomly assign students to the control and treatment groups. Therefore, we use a falsification test to generate confidence in our identification strategy as a solution. To ensure that our regression coefficient of the dummy $Treat_c * Post_t$ is not a result of any spurious relation but a causal effect of Home-Based Learning program, we conduct a falsification test for random simulation of treatment status. We randomly assign the treatment status to students across the control and the treatment groups. We replicate this process 100 times each for both of our outcome variables, $Maths$ and $Odia$. Ceteris paribus, this makes the association between our outcome variables and the treatment in our estimation random. Figs. 7(a) and 7(b) plots the t-statistics obtained from this analysis for our outcomes in comparison to the 90% confidence limits to illustrate the proportion of times we do not find significant effects.[27]

As is evident from Fig. 7(a), we do not find a statistically significant impact on the $Maths$ learning level 94 out of 100 times. Similarly, we do not find statistically significant impacts on the $Odia$ learning levels 91 out of 100 times. To be precise, we are unable to reject the null that the Home-Based Learning treatment effect is equal to zero 6 (9) out of 100 times for $Maths$ ($Odia$). This falsification analysis reveals that repeated estimations with random assignments of Home-Based Learning to students over time do not produce significant results in majority of the simulations on our outcomes of interest. This is evidence that our original specification is not picking up a spurious relation.

9. Conclusion, discussion and limitations

In this paper, utilizing a difference-in-differences strategy, we investigate the effects of a Home-Based Learning intervention on learning outcomes of children during COVID-19 induced school closures in Odisha, India. We find evidence that a well-designed, low-cost technology based interventions can be useful to cushion the decline in student learning among those belonging to low-resource communities. The recent School Children's Online and Offline Learning (SCHOOL) survey revealed that only 8 percent of rural students had access to online education, and at least 37 percent stopped studying altogether as a result of school closures due to the pandemic in India. Moreover, less than half of the rural households have access to a smart phone, and even within those households, students do not get a priority in using them. To promote continuity in learning, the Home-Based Learning intervention overcomes these limitations by providing remote instructions via phone calls and simple text messages, which can be accessed using any basic feature phone.

We find that exposure to the Home-Based Learning program results in improvements in student outcomes, measured in form of endline test scores in $Maths$ and $Odia$. The improvement in scores is consistent across different specifications, alternative estimation techniques, and sub-samples, thus providing robust evidence that the intervention resulted in better learning outcomes in the treatment group during the time when schools were shut, relative to students in the control group who were not a part of the program. Our paper joins other recent papers in testing the effect of low-cost phone based learning bringing clear evidence to this important topic (Ahluwalia et al., 2023; Angrist, Bergman, Evans et al., 2020; Angrist, Bergman and Matsheng, 2020; Crawfurd et al., 2021; Hassan et al., 2021; Radhakrishnan et al., 2021; Rodriguez-Segura & Schueler, 2022).[28]

Our results and the improvements in student learning levels ought to be appreciated in the broader context of pandemic-led school shutdowns (which have been the longest in the world) and lock-down movement restrictions (which were among the most severe in the world) in India. The learning deficit arising from being disconnected from an educational environment can have pernicious long term effects, ranging from immediate dropouts to spilling into poor health choices and lower social mobility for many first generation learners. Our results speak to the importance of such low-cost interventions having the capability to supplement education in a post-pandemic world wherein hybrid formats of teaching–learning are expected to become commonplace.

While we attempt to tackle an important problem by specifically targeting disadvantaged students during a pandemic, it is also essential to point out the limitations of our study. First, given that we were restricted due to the nation-wide lockdown, much of the intervention was conducted remotely (similar to other studies during the time), which resulted in differential attrition in the control and treatment groups. While we conduct several robustness checks using Lee (2009) bounds, propensity score matching estimation and inverse propensity weighting estimators, we acknowledge that we are unable to completely solve for this issue of attrition that took place as the program was remote and several students could not be contacted due to successive lockdowns and a devastating second wave of the pandemic in India. Second, though the intervention included and encouraged parental involvement by design, it does not measure the extent of their involvement and thus the analysis does not disentangle the effects that may be arising due to parental efforts in teaching their kids. However, while we cannot rule out the involvement of parents potentially driving the results, given the low literacy levels of parents in the sample, we believe their involvement in teaching at home is limited, especially for Maths.[29][30] Third, we recognize that we are unable to provide information on potentially mediating factors, such as students' interest in mathematics and language learning.

Like any turnaround after a crisis, there is no single way to ensure that students return to schools and are able to catch up with their peers with access to better resources. As such, teacher-assigned homework and parental involvement have been traditionally considered two important sources of learning in offline classes. The Home-Based Learning program tries to integrate both these aspects with the aim to keep students engaged in learning. Our results are in line with

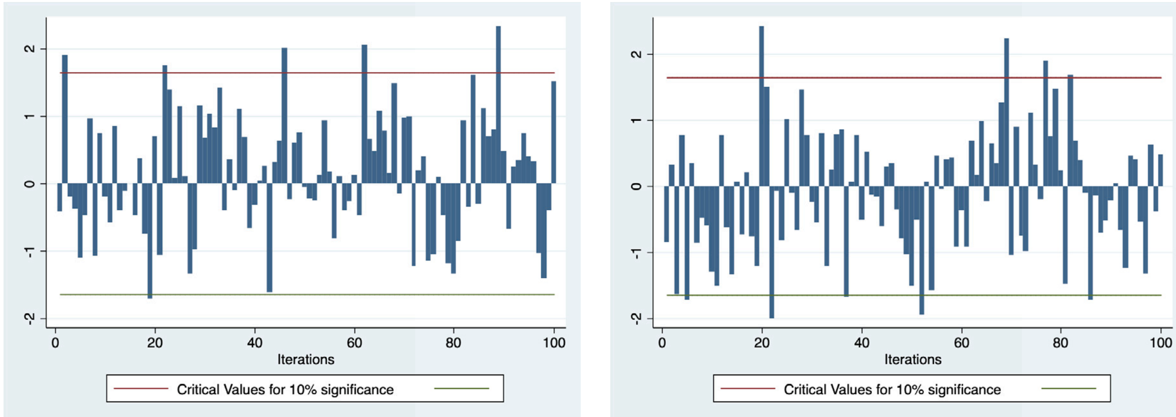
[26] The IPW estimator requires fewer functions to be estimated nonparametrically than other matching estimators.

[27] Note that since the 95% and 99% confidence intervals are larger than that for 90%, we would have even fewer instances of significant effects at the 5% and 1% level of significance.

[28] The approximate cost per student was approximately 400 INR (USD 5) for the entire duration of the intervention.

[29] Parental involvement is typically a greater worry in cases where the parents directly help their children answer questions in such phone-based assessments (Ahluwalia et al., 2023).

[30] We also provide some anecdotes from the parents in Appendix A.2. of the paper.



(a) Mathematics (b) Odia

Fig. 7. t-statistics distribution from 100 iterations with random assignment of treatment.

recent evidence by Ahluwalia et al. (2023), Kumar et al. (2022) and Rodriguez-Segura and Schueler (2022) that establish the reliability and validity of phone based assessments and recommend effectively integrating them in teaching–learning pedagogies.[31] That said, a drawback of such phone based programs is that it cannot be used to deliver complex academic modules for senior classes. It also does not provide any visual stimulus, which typically makes learning easier and increases retention power. While the program cannot (and ought not to) be seen as a substitute for regular schools or internet-based lessons, governments may consider investing in providing such safety nets especially for students with limited access to a smartphone or a school and wish to remain engaged with learning in the face of sudden shocks.[32] Allowing an opportunity to supplement morning school with an evening revision through a program like Home-Based Learning may reduce dependency on costly private tuition for the poor.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We are grateful to the Editor, Mauricio Romero, and two anonymous reviewers for their valuable advice and very helpful suggestions. We also thank Abhijeet Singh, Karthik Muralidharan, Alex Eble, Nishith Prakash, Ian K. McDonough, and Arnab Mukerji for their valuable insights and guidance. Any remaining errors are our own.

[31] Considering that the costs of verifying learning outcomes are prohibitively high, phone-based assessments are proving to be economical and prompt.

[32] With mobile learning proving an in-built flexibility that the under-resourced households value highly, the government ought to continue to invest in networks and provision of these devices to the under-resourced communities.

Funding statement

The authors have no funding to disclose for this research. ThinkZone would like to acknowledge the collaborative support received from MoSchool Abhiyan Parichalana Sangathan (MSAPS), a Government of Odisha Initiative under the School & Mass Education Department.

Appendix

A.1. Activity based learning content

Examples of the activity based content for basic number recognition and arithmetic operations and Odia are as follows:

Today, you will help your child learn to count. Take four spoons. Keep one spoon and say one loudly. Then keep another spoon and say two. Like this, by using spoon, help your child count till number four. Later, say to a number between 1 to 4 and ask your child to keep those many spoons.

Today, you will help your child learn to recognize letters. Take old newspapers. Cut five letters you want your child to learn today from the newspapers. Keep these letters in front of the child and read them out loudly. Ask the child to read the letters with you. Later, name one letter you taught, and ask the child to find that letter from the newspaper.

A.2. Parent's experiences

Mr. Rajesh from Kendrapada: *'I could not afford a smartphone last year and was worried that Sulagna would be unable to learn anything as her school was closed'. 'I was apprehensive if I would be able to support Sulagna's learning in the home-based learning program. But the activities shared on my small phone were so simple and engaging that I started enjoying practicing the same with Sulagna. In one of the activities, she had to note down the names of all her family members and what they like to eat. She made a list with so much enthusiasm, and I helped her write correct spellings for these names in the list. In another activity, she had to identify shapes using daily objects. I will continue practicing these activities at home with Sulagna even when school reopens as I feel proud that I can contribute to her learning in a way she can enjoy'.*

Mrs. Priyadarshini from Bhadrak: Priyadarshini's daughter Akankhya who has been enrolled in the nearby government school is in second grade now but has not gone to school for nearly two years. Despite this, she is happy to see Akankhya's progress during the pandemic. *'The home-based learning program has been beneficial as I receive simple daily activities over the phone, and I love practicing the same with Akankhya. The instructor calls once a week to check if I am struggling to understand the activity. The toll-free number facility has been useful as I can call on it to understand the learning activities in detail'* (see Figs. 8 and 9).

A.3. Fig. 8

Math: The baseline exam will be taken over the phone. The teacher will ask the questions to the student one after another and put marks (right/wrong) as per the answer given by the student.

Number	Addition	Subtraction	Multiplication	Division
Which number comes before 10?	5+7 =	7 – 3 =	9 × 3 =	9 ÷ 3 =
Which number comes after 19?	1+8 =	9 – 2 =	7 × 9 =	8 ÷ 2 =
What is the place value of 2 in 27?	23+18 =	33 –17 =	23 × 6 =	75 ÷ 5
What is the expanded form of 49?	20+33 =	44 – 31 =	39 × 8 =	36 ÷ 4 =

Fig. 8. Sample exam for measuring baseline and endline learning levels in mathematics.

A.4. Fig. 9

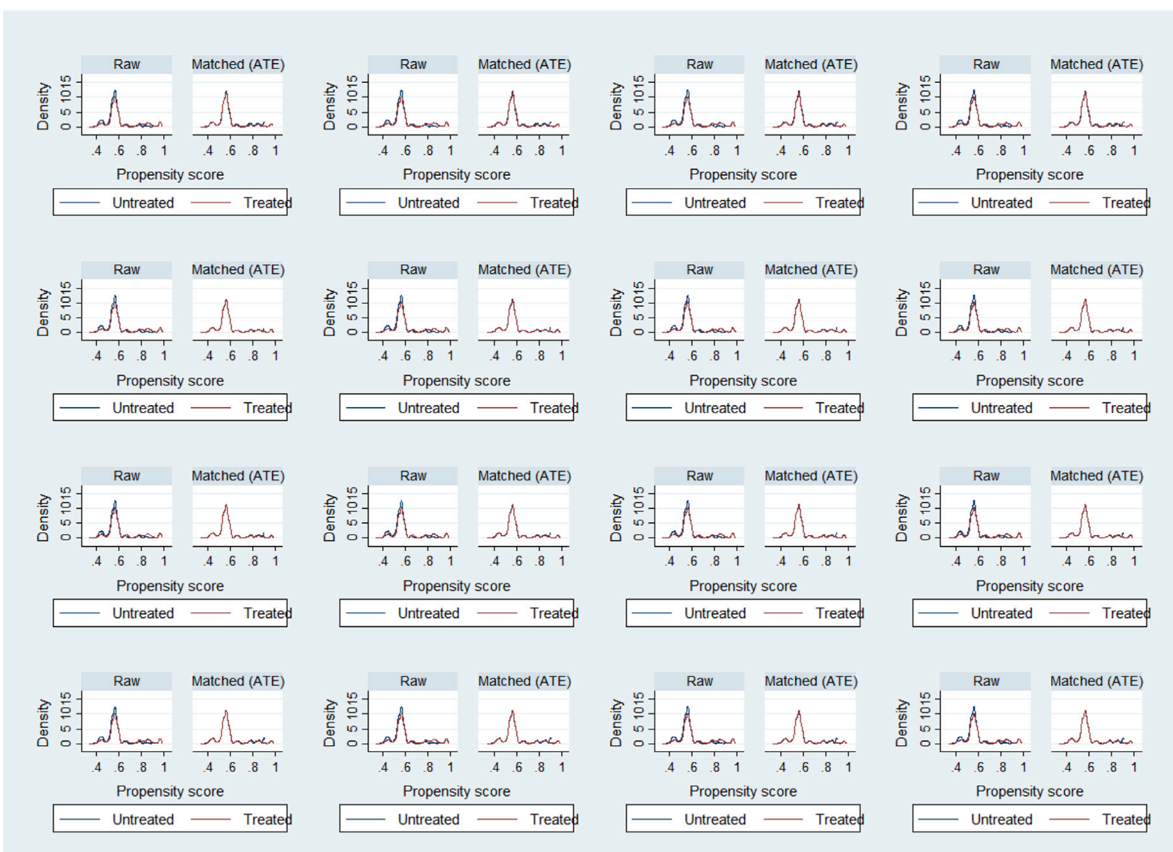


Fig. 9. Distribution of propensity scores, before and after matching. The first row graphs correspond to kernel matching, in the following order (L to R)- Score for Maths, Score for Odia, Improvement for Maths and Improvement for Odia. The second row corresponds to Nearest Neighbor (N = 1) matching, the third corresponds to N = 3 and the last row corresponds to N = 5 matching estimation.

References

Afridi, F., Barooah, B., & Somanathan, R. (2020). Improving learning outcomes through information provision: Experimental evidence from Indian villages. *Journal of Development Economics*, *146*, Article 102276.

Agragami, S. (2020). *The pandemic year: How 2020 unfolded in Odisha*. The New Indian Express, 31 December 2020. Available: https://www.newindianexpress.com/states/odisha/2020/dec/31/thepandemic-year-how-2020-unfolded-in-odisha-2243522.html [Last accessed: 1 July 2021].

Ahluwalia, R., Chopra, V., Gupta, S., Gupta, N., Panjwani, A., Satyam, K., et al. (2023). *Phone-based assessment data: Triangulating schools' learning outcomes*. Ideas for India.

Aker, J. C., & Ksoll, C. (2020). Can ABC lead to sustained 123? The medium-term effects of a technology-enhanced adult education program. *Economic Development and Cultural Change*, *68*(3), 1081–1102.

Aker, J. C., Ksoll, C., & Lybbert, T. J. (2012). Can mobile phones improve learning? Evidence from a field experiment in Niger. *American Economic Journal: Applied Economics*, *4*(4), 94–120.

Angrist, N., Bergman, P., Evans, D. K., Hares, S., Jukes, M. C., & Letsomo, T. (2020). Practical lessons for phone-based assessments of learning. *BMJ Global Health*, *5*(7), Article e003030.

Angrist, N., Bergman, P., & Matsheng, M. (2020). School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic.

ASER (2018). *Annual status of education report, 2018*: *Technical report*, Pratham.

ASER (2020). *Annual status of education report, 2020*: *Technical report*, Pratham.

Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A., & Geven, K. (2020). *Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates*. The World Bank.

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, *193*, Article 104345.

Bando, R., Gallego, F., Gertler, P., & Fonseca, D. R. (2017). Books or laptops? The effect of shifting from printed to digital delivery of educational content on learning. *Economics of Education Review*, *61*, 162–173.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., et al. (2016). *Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India*: *Technical report*, National Bureau of Economic Research.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *Quarterly Journal of Economics*, *122*(3), 1235–1264.

Barik, S. (2020). *Now Odisha turns to radio for classes*. The Hindu, 27 September 2020. Available: https://www.thehindu.com/news/national/other-states/now-odisha-turns-to-radio-for-classes/article32707975.ece [Last accessed: 1 July 2021].

Barrera-Osorio, F., Gertler, P., Nakajima, N., & Patrinos, H. (2020). *Promoting parental involvement in schools: Evidence from two randomized experiments*: *Technical report*, National Bureau of Economic Research.

Bergman, P., & Chan, E. W. (2021). Leveraging parents through low-cost technology the impact of high-frequency information on student achievement. *Journal of Human Resources*, *56*(1), 125–158.

Berlinski, S., Busso, M., Dinkelman, T., & Martínez, C. (2021). *Reducing parent-school information gaps and improving education outcomes: Evidence from high-frequency text messages*: *Technical report*, National Bureau of Economic Research.

Berry, J., Kannan, H., Mukherji, S., & Shotland, M. (2020). Failure of frequent assessment: An evaluation of India's continuous and comprehensive evaluation program. *Journal of Development Economics*, *143*, Article 102406.

Carlana, M., & La Ferrara, E. (2021). *Apart but connected: Online tutoring and student outcomes during the COVID-19 pandemic*: *edworkingpaper no. 21-350*, Annenberg Institute for School Reform At Brown University.

Coffey, D., Hathi, P., Khalid, N., Khurana, N., & Thorat, A. (2021). *Phone survey methodology for social and economic research in India*. Ideas for India.

Crawfurd, L., Evans, D. K., Hares, S., Sandefur, J., et al. (2021). *Teaching and testing by phone in a pandemic*: *Technical report*, Center for Global Development.

Doss, C., Fahle, E. M., Loeb, S., & York, B. N. (2019). More than just a nudge supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, *54*(3), 567–603.

Economist, T. (2021). *Closing the world's schools caused children great harm*. The Economist, 24 June 2021. Available: https://www.economist.com/leaders/2021/06/24/closing-the-worlds-schools-caused-children-great-harm [Last accessed: 29 June 2021].

Escueta, M., Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). Upgrading education with technology: Insights from experimental research. *Journal of Economic Literature*, *58*(4), 897–996.

Frölich, M. (2004). Finite-sample properties of propensity-score matching and weighting estimators. *The Review of Economics and Statistics*, *86*(1), 77–90.

Government of Odisha (2018). *Odisha economic survey report 2017-18*. Available: http://orienvis.nic.in/indexx.aspx?langid=1&slid=735&mid=2&sublinkid=236. Last accessed 21 January 2023.

Gurudu, D. (2021). *Rural internet connectivity in India: Gaps and challenges*. The New Learn, 2021. Available: https://www.thenewleam.com/2021/05/rural-internet-connectivity-in-india-gaps-and-challenges/ [Last accessed: 4 Jan 2023].

Hassan, H., Islam, A., Siddique, A., & Wang, L. C. (2021). "Low-tech" distance education: A remedy of the learning loss due to the COVID-19 pandemic.

Hindu, T. (2021). *COVID-19-induced school closures affected 25 crore Indian children: UNICEF study*. The Hindu, 3 March 2021. Available: https://www.thehindu.com/news/national/closure-of-15-million-schools-due-to-covid-19-impacted-247-million-children-in-india-unicef-study/article33981143.ece [Last accessed: 29 June 2021].

HindustanTimes (2021). *Students of Class 1 to 8 may be promoted without any exam: Odisha Minister*. HindustanTimes, 14 March 2021. Available: https://www.hindustantimes.com/education/news/students-of-class-1-to-8-may-be-promoted-without-any-exam-odisha-minister-101615700085808.html [Last accessed: 1 July 2021].

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), 1161–1189.

Hossain, M. (2020). COVID-19: Online learning inequality in four developing countries.

ILO (2018). *Guide on measuring decent jobs for youth monitoring, evaluation and learning in labour market programmes*. Geneva: ILO.

Kawoosa, V. M. (2020). *Connectivity gets better but parts of India still logged out*. Hindustan Times, 2020. Available: https://www.hindustantimes.com/india-news/connectivity-gets-better-but-parts-of-india-still-logged-out/story-VSqXriMdGUudWb7eBcWzjN.html [Last accessed: 4 Jan 2023].

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253.

Ksoll, C., Aker, J. C., Miller, D., Perez, K., & Smalley, S. L. (2015). *Learning without teachers? Evidence from a randomized experiment of a mobile phone-based adult education program in Los Angeles*: *Technical report*, CGD Working Paper.

Kumar, S., Gupta, N., Ahluwalia, R., & Gupta, S. (2022). *Can phone assessments generate reliable student learning data?* Central Square Foundation.

Lakshminarayana, R., Eble, A., Bhakta, P., Frost, C., Boone, P., Elbourne, D., et al. (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), Article e65775.

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

Lichand, G., & Christen, J. (2021). *Behavioral nudges prevent student dropouts in the pandemic*: *Working Paper (363)*, University of Zurich, Department of Economics.

Mohanty, D. (2020a). *No immediate possibility of opening schools in Odisha, says minister*. HindustanTimes, 15 August 2020. Available: https://www.hindustantimes.com/india-news/no-immediate-possibility-of-opening-schools-in-odisha-says-minister/story-s7cdA5A6YrfcUyZyTGVHfL.html [Last accessed: 1 July 2021].

Mohanty, D. (2020b). *Odisha class 10 and 12 exams to get delayed, exam only after 3 months of classroom education*. HindustanTimes, 18 November 2020. Available: https://www.hindustantimes.com/education/odisha-class-10-and-12-exams-to-get-delayed-exam-only-after-3-months-of-classroom-education/story-VajXaPFDSJNQUK8kx3Hh8J.html [Last accessed: 1 July 2021].

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, *109*(4), 1426–1460.

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, *119*(1), 39–77.

Naik, G., Chitre, C., Bhalla, M., & Rajan, J. (2020). Impact of use of technology on student learning outcomes: Evidence from a large-scale experiment in India. *World Development*, *127*, Article 104736.

OESPA (2019). *Status of elementary and secondary education in Odisha 2018-19, Odisha school education programme authority*. Management Information System, OESPA, 2019. Available: //http://osepa.odisha.gov.in/webadmin/pdf_link/UDISE%202018-201_07_04_37pm9d4773d0dfc121c936df3252a79808e0.pdf [Last accessed: 26 Feb 2023].

Radhakrishnan, K., Sabarwal, S., Sharma, U., Cullen, C., Crossley, C., Letsomo, T., et al. (2021). Remote learning.

Rodriguez-Segura, D. (2021). EdTech in developing countries: A review of the evidence. In *The world bank research observer*.

Rodriguez-Segura, D., & Schueler, B. E. (2022). Can learning be measured by phone? Evidence from Kenya. *Economics of Education Review*, *90*, Article 102309.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, *19*(1), 4–60.

Sabates, R., Carter, E., & Stern, J. M. (2021). Using educational transitions to estimate learning loss due to COVID-19 school closures: The case of Complementary Basic Education in Ghana. *International Journal of Educational Development*, *82*, Article 102377.

Schueler, B. E., & Rodriguez-Segura, D. (2022). A cautionary tale of tutoring hard-to-reach students in Kenya. *Journal of Research on Educational Effectiveness*, 1–31.

Seid, Y. (2016). Does learning in mother tongue matter? Evidence from a natural experiment in Ethiopia. *Economics of Education Review*, *55*, 21–38.

Silverman, B. (1986). *Statistics and applied probability*, *Density estimation for statistics and data analysis*.

Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, *94*, 252–275.

Tauchmann, H. (2014). Lee (2009) treatment-effect bounds for nonrandom sample selection. *The Stata Journal*, *14*(4), 884–894.

UNESCO (2021). *Education: From disruption to recovery*. UNESCO: COVID-19 impact on education, 2021. Available: https://en.unesco.org/covid19/educationresponse [Last accessed: 29 June 2021].

UNICEF (2021). *COVID-19: Schools for more than 168 million children globally have been completely closed for almost a full year*. UNICEF, Press Release, 3 March 2021. Available: https://www.unicef.org/india/press-releases/covid-19-schools-more-168-million-children-globally-have-been-completely-closed [Last accessed: 29 June 2021].

UNICEF, et al. (2020). *How many children and young people have internet access at home?: estimating digital connectivity during the COVID-19 pandemic*: *Technical report*, UNICEF.

Vagh, S. B. (2012). Validating the ASER testing tools: Comparisons with reading fluency measures and the Read India measures. Unpublished Report. Retrieved July, 30:2018.

Yanguas, M. L. (2020). Technology and educational choices: Evidence from a one-laptop-per-child program. *Economics of Education Review*, *76*, Article 101984.

York, B. N., Loeb, S., & Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, *54*(3), 537–566.