Not Too Little, But Too Late: Improving Post-Primary Learning Outcomes in India*

Gaurav Chiplunkar
University of Virginia

Diva Dhar
University of Oxford

Radhika Nagesh
UNICEF

March 2022

Abstract

We evaluate an innovative remedial education program targeted at post-primary school students and aimed at improving their learning outcomes ahead of an important standardized national exam in India. While the program leads to around 0.5σ gains in basic reading and math scores, it fails to translate into better scores on the national exam. We show that this is not because the program crowds out students' time and effort in studying grade-level material, but because these Grade 9 students lag far behind (i.e., have Grade 3 competencies). We conclude that in settings with severe learning deficiencies in advanced grades, such programs can improve basic learning, but may not be sufficient to overcome accumulated academic deficiencies at grade-level.

JEL Codes: I21, I26, I28, J24

Keywords: Post-primary education, Secondary School, Remedial Education, India

*We thank Prashant Bharadwaj, Alejandro Ganimian, Paul Glewwe, Clare Leaver, Anandi Mani, Isaac Mbiti, Karthik Muralidharan, Dipa Nag Chowdhury, Abhijeet Singh as well as participants at the RISE, PacDev and MWIEDC conferences and BSG and CSAE seminars for helpful comments and suggestions. We would also like to thank Avanti Fellows, especially Akshay Saxena and Deepak Kamble for their collaboration, and MacArthur Foundation for their financial support. Contact information: Chiplunkar *(corresponding author)*: Darden School of Business, University of Virginia; Email: ChiplunkarG@darden.virginia.edu; Dhar: Blavatnik School of Government, University of Oxford; Email: diva.dhar@bsg.ox.ac.uk; Nagesh: UNICEF Innocenti Office; Email: radhika.n89@gmail.com. All views expressed are those of the authors and not of any of the institutions with which they are affiliated.

1 Introduction

Investment in education, to improve human capital development and boost "inclusive growth", has been at the cornerstone of development policy across the world. However, despite an increase in student enrollments worldwide, learning outcomes have stagnated, especially in low- and middle-income countries.[1] Moreover, these learning deficiencies continue to accumulate over time, as students fall behind academically and progress through the school system, absorbing and learning very little from their grade level classroom instruction, which remains above their learning levels (Banerjee and Duflo, 2011). For example, 61 percent of adolescents across the globe are unable to achieve minimum proficiency levels in Math and Reading upon completing lower secondary school (UNESCO, 2017; World Bank, 2018). This has further implications on their subsequent education choices, early employment outcomes, and long-run socio-economic status and mobility more broadly (Das, Singh and Chang, 2022). While a large literature provides various explanations for the continued deficiencies in learning levels at the primary level[2], there remains little evidence on how policies and interventions can impact these outcomes at the post-primary or secondary school level. Understanding the effectiveness of such programs can also be of relevance to public policy, given that most countries across the world spend over a third of their education budget on secondary education (World Bank, 2020).

This paper aims to fill this gap by reporting on a randomized evaluation of one such technology-aided, after-school instruction program at the post-primary school level (Grade 9) in India. With around half a billion children below the age of 18, the Indian context is no exception to the above global trends. India has achieved near-universal school enrollment rates in primary education, 85 percent enrollment in secondary education (Pratham, 2016), and has made significant improvements in areas of school access, infrastructure, pupil-teacher ratios and teacher salaries over the last decade (Muralidharan, 2013). However, learning at the post-primary level continues to remain poor. For example, in a nationally representative sample of Grade 8 students, more than half (55

[1]See Angrist, Djankov, Goldberg and Patrinos (2021); Glewwe and Muralidharan (2016); Murnane and Ganimian (2016).

[2]See see Kremer, Brannen and Glennerster (2013), Glewwe and Muralidharan (2016) and Evans and Popova (2016) for extensive reviews.

percent to be exact) could not read a simple English sentence and 57 percent could not do a simple mathematical operation (Pratham, 2016). In our sample of Grade 9 students as well (see Figure 1), around 65 percent of students could not read a story and around half the students could not complete basic multiplication and division. These learning deficiencies become particularly salient in the presence of mandatory national exams (like Grade 10 in India), which play an important role in determining further educational choices as well as subsequent labor market outcomes.

It is in this context that we evaluate the impact of an innovative after-school instruction program developed by Avanti, and implemented with Grade 9 students across 24 public schools in Chennai, India. The program incorporates the use of technology into a remedial instruction program, where trained facilitators make effective use of technology aids (video recordings, presentations, worksheets etc.) along with peer learning and counseling, to enhance students' cognitive and non-cognitive development. The incorporation of technology aids, in an otherwise standard remedial program, aims to improve the quality of education for marginalized students at minimal costs to the schools. We examine whether the program improves basic learning outcomes (as measured by ASER test scores), as well as how it translates into improvements in test scores on a standardized, mandatory national exam one year later.

We find that Avanti's program leads to large gains in basic reading and math ASER scores (0.47σ and 0.52σ respectively). These effects are comparable to similar remedial education programs in India like the Teaching at the Right Level evaluation (Banerjee et al., 2016) and the MindSpark program (Muralidharan, Singh and Ganimian, 2019). However, we find that these large gains in foundational learning outcomes do not translate into improvements in test scores in the Grade 10 national exam one year later. We further show that this is not because the program substituted or crowded out time/effort that students could put in for grade-level material (such as absenteeism in school, parental support, attending private tuition classes, self-study, etc.), but because the accumulated learning deficiencies at grade-level seem too vast to bridge. On average, these Grade 9 students have reading and math skills that are required in Grades 2 or 3 (see Figure 1).

Put together, we conclude that in contexts where there are substantial learning deficiencies at grade level, carefully designed remedial education interventions can bring large gains in basic learning levels. However, at the post-primary level, it might be too late to translate these large gains in foundational skills into meaningful progress on school completion and further advanced human capital accumulation. This highlights the importance of policy debates in developing countries to address the mismatch between student learning and grade-level requirements (Muralidharan et al., 2019). Furthermore, it reiterates the importance of the timing of implementing these programs to mitigate learning deficiencies at the primary school level itself instead of letting them accumulate across grades.

Our paper contributes directly to a growing body of research that has evaluated the impact of pedagogical interventions to improve learning outcomes and test scores. These include Teaching at the Right Level (Banerjee et al., 2007, 2016), computer assisted learning (Linden et al., 2003; Lai et al., 2015) including adaptive technology-aided instruction (Muralidharan, Singh and Ganimian, 2019), curriculum-based expert-led instructional videos (Beg et al., 2020), low-technology interventions such as SMSes and direct phone calls (Angrist et al., 2020), model schools (Kumar, 2020), activity based learning (de Barros et al., 2020) and performance-based incentives for teachers (Muralidharan and Sundararaman, 2011). While this literature has primarily focused on improving outcomes in primary schooling, we evaluate the impact of a similar remedial education program that is implemented at the post-primary level.

Our paper also complements evidence on "what works" in improving learning outcomes in developing countries more generally, such as provision of information on school quality (Andrabi et al., 2015); student level scholarships and incentives (Blimpo, 2014; Kremer, Miguel and Thornton, 2009); improved access through free bicycles for girls (Muralidharan and Prakash, 2017); curricular simplification and pacing (Mbiti and Rodrigue-Segura, 2020; Pritchett and Beatty, 2012); public rankings of schools (Cilliers, Mbiti and Zeitlin, 2019); and bundled interventions such as combining school grants with teacher incentives (Mbiti et al., 2019).

The rest of the paper is organized as follows: Section 2 describes the empirical context and the experimental design. Section 3 describes the data collection and the sample. Section 4 discusses the empirical strategy and results. Section 5 offers a short conclusion.

2 Empirical Context

2.1 Avanti's Program

Avanti's approach combines trained facilitators, technology aids and parent counseling, along with an emphasis on peer learning and team work.[3] Avanti's pedagogy incorporates features from Eric Mazur's peer instruction and collaborative learning pedagogy (Fagen et al., 2009; Schell et al., 2013; Zhang et al., 2017) and other blended learning programs. Avanti has adapted this methodology for low-resource settings (common to educational settings in developing countries and in particular, India) for teaching Maths and Science in government secondary schools in India, using basic technology and infrastructure available or installed in schools, such as computers (without internet).

Conducted after-school for an hour for five days a week, Avanti classes run for 40 weeks in an academic year, led by trained facilitators rather than teachers. Facilitators are usually fresh graduates who are required to have a Bachelor's degree, preferably (but not necessarily) in STEM, and do not having teaching certification or qualifications.They are trained on facilitating classroom sessions which rely on peer learning, team work and student collaboration. The facilitators rely on pre-recorded video lectures, presentations, worksheets and other academic materials developed by Avanti for Math and Science in the regional language and context. In the study schools, the program materials and discussions are both in Tamil (the local native language), although a school's official medium of instruction may be English or Tamil. Finally, the program also helps provide occasional career counseling and guidance for students and parents. Table B1 provides a summary of the principal components of the Avanti program.

2.2 Experimental Design

The study was conducted in a sample of 24 government schools under the management of the Chennai Municipal Corporation (CMC) in the city of Chennai, India. Of the total 281 schools

[3]Peer learning for example, engages students during class through worksheets and multiple choice questions (ConcepTests), where they are organized into learning groups based on their academic achievement levels and collaboratively encouraged to solve the worksheets. Facilitators move from group to group, encouraging discussion and resolving queries when necessary.

under the CMC, data on Grade 9 students for 24 schools was provided by the Education Department of the Corporation of Chennai. Schools were selected based on the following eligibility criteria: (a) minimum of 45 students in Grade 9;[4] (b) maximum of 110 students;[5] (c) no previous experience or engagement with Avanti (to avoid contamination). The experimental design involved introducing the Avanti program in 12 randomly selected schools (henceforth, treatment schools) for the academic year 2016-17, leaving students from the remaining 12 schools to serve as a comparison group (control schools). The selection was done in an open paper lottery at the Chennai Municipal Corporation, under the observation of the research team to ensure that the selection was fair. Schools complied with the treatment selection and the program was introduced for all Grade 9 students in the treatment schools. From each sample school, 45 students were randomly sampled based on student registers at each school for a survey at the beginning of the academic year (henceforth, Baseline). However, out of 1080 students listed from student registers, 36 students could not be surveyed either due to school absence on the days the survey was taking place in their schools, or because they had changed or dropped out of that school between enrollment and the baseline survey. Our study sample therefore comprises of 1,044 students at baseline (97 percent of the listed students). During the year, some students dropped out or transferred to other schools resulting in a final (henceforth, Endline) sample of 841 students one year later (81 percent of the baseline sample).

3 Data

We implemented three sets of student surveys at baseline and endline. The first survey included questions on students' demographic and household information, academic performance, aspirations and expenditure on education (books, tuition etc.). Second, we measured learning outcomes at the start and end of the program, using the standard ASER testing tool for Reading and Mathematics developed and implemented by ASER. The ASER instrument has been widely adopted in India (and other countries) to assess students' mastery over foundational skills at different stages of

[4]Based on power calculations at 80 percent power and statistical significance at 5 percent level, assuming a 10 percent sample attrition rate between the baseline and endline.

[5]For logistical reasons, Avanti could not handle schools with more than 110 students.

education. It is also frequently used in evaluations, such as those of Pratham's Read India program (Banerjee et al., 2010). The ASER test uses a five-point grading system and covers up to Grade 2 level reading skills and up to Grade 4 level mathematics ability . The test is administered orally for each student, by trained enumerators, and lasts up to ten minutes, following standard administration protocols laid out by ASER.[6] In addition to the ASER scores, we also collected exam scores for all subjects in the national standardized Grade 10 examination in 2018, and matched them to students in the sample. Third, we conducted a short survey with students' parents to better understand parental involvement in, attitude towards, and support for their child's education, measured as cumulative category scores using a series of composite Likert scales. Lastly, we conducted a battery of tests to measure gender attitudes and non-cognitive skills such as critical thinking, communication, goal-setting, problem solving, grit, self-esteem and team work (West et al., 2014; Duckworth and Quinn, 2009). The survey tools were administered in the local language (Tamil).[7]

3.1 Sample Description

Panel A of Table 1 describes the characteristics of the 1,044 students in our baseline sample in Column (1). Around 41 percent of students in our sample are female, and students are on average 14-15 years of age (which is appropriate for Grade 9 students). Around 85 percent of students come from socially disadvantaged backgrounds (Scheduled Castes, Scheduled Tribes and Other Backward Castes). From Panel B, 20 percent of students have parents who are illiterate, and nearly 80 percent of households earn a monthly income less than Rs. 10,000 (PPP-adjusted USD 550). Columns (2) and (3) of Table 1 then report the averages separately for students in the control and treatment schools, while Column (4) reports the difference and its statistical significance at conventional levels. We see that there are no statistically significant differences between the demographics of students in the treatment and control schools and these differences are also very small in magnitude.

[6] Information on the modalities of conducting the ASER test and tools is retrieved from aser-centre.org/Survey/Basic/Pack/Sampling/History/p/54.html

[7] While we do not discuss the impact of our intervention on non-cognitive skills (such as communication, problem solving, gender attitudes, etc.), we do measure them and find no impact. Results are available upon request.

Panel C describes the cognitive and non-cognitive ability of students as well as the parental support they receive. As described before, we use the ASER tool to measure basic cognitive ability in reading and math (for details, see Appendix B). The reading scores, for example, indicate that students are able to read words and short sentences on average, but not paragraphs. Similarly, on basic math, they can do simple operations of addition/subtraction, but not more complicated ones such as multiplication/division. These scores indicate that an average Grade 9 student has only mastered skills usually taught in Grades 2 and 3, and are lagging far behind in the grade-level skills needed. This finding resonates with other studies in the Indian context that document the wide learning gap at the post-primary level (Pratham, 2016).

It is important to note from Column (4), that students in treatment schools perform better than control students at baseline on reading and math– on average, they are ahead by nearly one level in Reading and around half a level in Math, and receive less parental support and supervision. Since these characteristics are important outcome variables in our analysis, we analyze the impact of our intervention in multiple ways to take this into account, as outlined in the subsequent sections.

4 Results

We now turn to examining the effect of the intervention for an individual i in school s. We estimate the following regression specification:

$$Y_{isE} = \alpha + \beta T_s + \gamma_1 Y_{isB} + \delta_1 X_i + \delta_2 X_s + \varepsilon_{is}$$

where Y_{isE} and Y_{isB} are the endline and baseline outcome variables of an individual i, T_s is a binary indicator that takes the value 1 (0) for a treatment (control) school, X_i and X_s are time-invariant characteristics of the individual and school respectively. We use age, gender, religion, dummy for caste (SC/ST, OBC and others) for X_i and number of boys and girls, indicator of a co-education school and language of instruction for X_s. We estimate the above specification both without and with individual and school controls and report them in Panels A and B in each table respectively. Lastly, since the randomization was done with 24 schools, we wild-bootstrap cluster our standard errors at the school level, as suggested by Cameron, Gelbach and Miller (2008) for statistical inference with small clusters. This is reported in all tables as the p-val (OLS) value. Furthermore, we also report the p-value from a two-sided randomization inference test, denoted as p-val (RI) in all tables. This test, originally proposed by Fisher (1935) and developed by Heß (2017) and Young (2019), allows for statistical inference by comparing the realized treatment effect with multiple (250) placebo assignments. This procedure therefore has the advantage of providing inference with correct size, regardless of the sample and cluster size.

4.1 Impact on Basic Learning Outcomes

We begin by examining the effect of the treatment on basic learning outcomes, as measured by the ASER reading and math scores. Both the reading and math scores can take a value from 1 to 5 as described earlier. We examine the impact of our intervention in three ways as discussed below.

The first set of results are reported in Table 2, where Panel A reports the results without any individual and school characteristics, while Panel B controls for them. In Columns (1) and (2), we standardize the ASER scores to have mean zero and standard deviation 1 in the control group, so that the treatment effect can be interpreted in standard deviations. We make two observations before discussing the results. First, baseline and endline scores are strongly positively correlated and second, controlling for individual and school characteristics does not substantially alter the magnitude or statistical inference of the impact of the intervention. Turning to the impact of our intervention, from Panel B (our preferred specification), after controlling for baseline ASER scores and characteristics, students in treatment schools have a 0.47 and 0.52 standard deviations higher ASER reading and math scores respectively as compared to students in the control school at the endline. This is equivalent to a 18.5 percent and 11.6 percent gain in the reading and math scores respectively as compared to the control group students. As a benchmark, this impact compares favorably with similar programs evaluated in the Indian context using ASER, such as Teaching at the Right Level (Banerjee et al., 2016) and Mindspark programs (Muralidharan et al., 2019). Since the ASER scores are discrete and ordinal, we also use an alternate specification and report results from a rank-ordered logistic regression in Columns (3) and (4). The results and their statistical inference are qualitatively similar for both outcome variables.

Given that students in the treatment schools have a higher ASER score at baseline (Table 1), in Appendix Table A1, we report the difference in the endline cognitive level by comparing treatment and control students with the *same* cognitive level at baseline. For example, in Column (1), we restrict the sample of students (in both treatment and control schools) to only those who could read at least a word at baseline. We find that students in treatment schools have a 0.25 higher ASER reading score at endline as compared with their control school counterparts. Similarly in Column (2), treatment school students who could read a paragraph at baseline have a 0.22 higher ASER score at endline as compared to their control school counterparts who could also read a paragraph at baseline. Lastly in Column (3), we see that among students who could read a story at baseline, those in treatment schools have a 0.28 higher (though statistically insignificant at conventional values) reading score at endline as compared to their control school counterparts with the same baseline reading ability. Similarly, turning to Math scores in Columns (4)-(6), we find that students in the treatment schools have a higher math score at endline, as compared to their control school counterparts who start at the same baseline math level. Lastly, as expected, the largest gains accrue to the students who have weaker skills at baseline. Put together, we conclude from the above analysis that the Avanti program intervention robustly resulted in substantial improvements in basic reading and math cognition of students.

4.2 Impact on Standardized National Exam Score

It is mandatory for all students in India completing Grade 10 to write a standardized national exam– the All India Secondary School Examination (AISSE). Students are tested in five subjects, namely: Math, English, Science, Social Sciences and the local language (Tamil, in our case). The AISSE exam is an important milestone in a student's education because these scores are used for future admissions in higher education as well for early career job applications.

Given the large improvement in basic cognitive ability discussed in the previous section, we examine whether this gain translates into an improvement in the AISSE scores. For the students in our sample, we obtain the list of their AISSE scores for all the five subjects in all of our 24 schools. Along with the list of individual and school characteristics, we now also control for baseline cognitive ability in the form of the baseline ASER score . We report the results without and with the ASER controls in Panels A and B of Table 3 respectively. Firstly, we see that baseline ASER score is strongly positively correlated with the AISSE exam scores. However, the treatment does not have any significant impact on these exam scores. The estimated magnitudes are very small ($\pm$1-2 percentage points) across all subjects, and statistically insignificant at conventional levels.

4.3 Discussion

The results so far indicate that while the intervention did significantly improve basic cognitive skills, it failed to translate into gains in standardized national exam scores for those students one year later. The natural explanation to rationalize this observation is that the learning deficiencies at the post-primary level are so severe (Figure 1) that despite significant improvements in basic reading and math skills, students are not able to catch up to the grade-level requirements. However, since the Avanti program required students to stay back in school for an hour for five days in a week, it could have crowded out time and efforts from students which they could have allocated towards studying grade-level material.

We therefore examine the impact of our intervention on various measures of this time and effort, such as absenteeism in school, parental support in studying, participation in private tuition classes, and time spent on self-study after school. As reported in Table 4, we find no effect of our intervention on any of these activities. Moreover, the estimated coefficients are small in magnitude, and in some cases go in the opposite direction that would be needed to explain away our results through these channels. For example, as reported in Column (1), we find that there is no effect of our intervention on absenteeism in school, as measured by the number of days in the previous week that a student was absent from school. From Column (2), we find no impact of the intervention on an index of parental support (that captures how involved parents were with school events, homework, participation in school activities, etc.).[8] If anything, the estimated

[8]The questions used to construct the index of parental support is described in Apprndix B. Basically, it aggregates responses on a Likert scale from 1-5 across four indicators of how often parents help their child with school work,

coefficients indicate a small reduction in absenteeism (by around 5 percent) and an increase in parental support (by less than 2 percent). Similarly, in Columns (3)-(6), we find no effect of the intervention on either the probability of attending private tuition classes or studying at home (Columns 3 and 5 respectively), or the amount of time spent on either activity (Columns 4 and 6 respectively). Put together, we find no evidence that our intervention crowded out effort and time that students could allocate to studying grade-level material. We thus conclude that while the intervention improved basic learning, the gains were not large enough to mitigate the substantial learning deficiencies acquired by the post-primary level.

4.4 Attrition

As mentioned previously, even though we had a sample of 1044 students at baseline, we were only able to measure outcomes for 841 students at the end of the academic year (an attrition of 20.3 percent). This attrition was caused both due to students dropping out of school as well as transfers to other schools, for example, due to family migration. In Section C in the Appendix, we test for differential attrition between control and treatment groups and check whether this affects our estimates. Specifically, we do so in two ways: first, we examine whether student and parent characteristics, cognitive and non-cognitive ability, and parental support are differentially correlated with either dropouts (Appendix Table C1), or the propensity to take the ASER tests (Appendix Table C2) across treatment and control schools. We find no evidence of differential attrition. Second, to account for potential endogenous attrition from the sample, we estimate Lee bounds on the treatment effects (Lee, 2009). These bounded estimates are consistent with our main analysis. For example, Appendix Table C3 shows that the attrition-adjusted lower bound on the impact of our intervention is 0.17σ for Reading and 0.16σ for Maths, while the bounds on the impact of the intervention on Class 10 exam scores across all subjects are very small in magnitude (± 1 percentage points). This increases our confidence in the robustness of our results after taking into account the attrition of students from our sample.

homework, attending events and talking about school activities. As reported in Table A2 in the Appendix, the intervention has no impact on any of these indicators. The estimated coefficients are very small in magnitude and statistically insignificant at conventional levels.

5 Conclusion

This paper reports on the results of an evaluation of a technology-based remedial education program in India, which was specifically targeted towards post-primary school students in Grade 9. We find significant learning deficiencies at baseline, with students having reading and math levels at Grade 2 or 3 on average. We find that the daily-run program, led by facilitators (rather than teachers) with technological inputs and academic materials, was successful in improving basic learning outcomes of the students. The effect sizes are large, and the program cost is reasonable at USD 50 per student per annum, given the program intensity and the powerful combination of education technology with in-person facilitation and teamwork. However, an important lesson from this study is that these gains fail to translate into any improvement on a national standardized Grade 10 exam, indicating that such programs targeted at the post-primary level may not be sufficient to overcome the accumulated learning deficiencies at grade-level. As Pritchett and Beatty (2012) note, the regular grade-level curriculum may be excessively difficult for learners in these settings, who have continued to accumulate deficiencies over many years and still to need to build on basic skills. Moreover, our findings indicate the urgency of addressing learning levels earlier in primary school itself, and rethinking post-primary educational initiatives in the context of high learning gaps. This also reinforces the importance of models of 'Teaching at the Right Level' (Banerjee et al., 2016), which is more cost-effective and helps ensure that learning happens effectively based on the child's needs rather than the dictates of the curriculum or the schooling system. Lastly, we hope to inform the debate around the targeting and implementation of India's (newly drafted) National Education Policy (2020), which highlights the importance of foundational numeracy and literacy, and the role of education technology for achieving educational goals.

References

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Delivering Education: A Pragmatic Framework for Improving Education in Low-Income Countries," *World Bank Policy Research*, 2015.

Angrist, Noam, Peter Bergman, Caton Brewster, and Moitshepi Matsheng, "Stemming Learning Loss During the Pandemic: A Rapid Randomized Trial of a Low-Tech Intervention in Botswana," *Available at SSRN 3663098*, 2020.

_ , Simeon Djankov, Pinelopi K Goldberg, and Harry A Patrinos, "Measuring Human Capital using Global Learning Data," *Nature*, 2021, *592* (7854), 403–408.

Banerjee, Abhijit V. and Esther Duflo, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs, 2011.

_ , Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani, "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," *American Economic Journal: Economic Policy*, 2010, *2* (1), 1–30.

_ , _ , James Berry, Esther Duflo, Harini Kannan, Shobhini Mukherji, Marc Shotland, and Michael Walton, "Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of Teaching at the Right Level in India," Technical Report, National Bureau of Economic Research 2016.

_ , Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying Education: Evidence from Two Randomized Experiments in India," *The Quarterly Journal of Economics*, 2007, *122* (3), 1235–1264.

Beg, Sabrin, Adrienne M. Lucas, Waqas Halim, and Umar Saif, "Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not," 2020.

Blimpo, Moussa P., "Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin," *American Economic Journal: Applied Economics*, 2014, *6*, 90–109.

Cameron, Colin A., Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Cilliers, Jacobus, Isaac Mbiti, and Andrew Zeitlin, "Can Public Rankings Improve School Performance? Evidence from a Nationwide Reform in Tanzania," *IZA Institute of Labor Economics*, 2019, *12172.*

Das, Jishnu, Abhijeet Singh, and Andres Yi Chang, "Test Scores and Educational Opportunities: Panel Evidence from Five Developing Countries," *Journal of Public Economics*, 2022, *206.*

de Barros, Andreas, Johanna Fajardo-Gonzalez, Paul Glewwe, and Ashwini Sankar, "Learning by Doing? Experimental Evidence on Activity-based Instruction in India," *Journal of Development Economics*, 2020.

Duckworth, Angela and Patrick Quinn, "Development and Validation of the Short Grit Scale (GritS)," *Journal of Personality Assessment*, 2009, *91*, 166–174.

Evans, David K. and Anna Popova, "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews," *The World Bank Research Observer*, 2016, pp. 242–270.

Fagen, Adam P., Catherine H. Crouch, and Eric Mazur, "Peer Instruction: Results from a Range of Classrooms," *The Review of Economics and Statistics*, 2009, *91* (3), 437–456.

Fisher, Ronald A, "The Design of Experiments," 1935.

Glewwe, Paul and Karthik Muralidharan, "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications," *Handbook of the Economics of Education*, 2016, *5*, 653–743.

Heß, Simon, "Randomization Inference with Stata: A Guide and Software," *The Stata Journal*, 2017, *17* (3), 630–651.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, "The Challenge of Education and Learning in the Developing World," *Science*, 2013, *340* (6130), 297–300.

_ , **Edward Miguel, and Rebecca Thornton**, "Incentives to Learn," *The Physics Teacher*, 2009, *40*, 206–209.

Kumar, Naveen, "Public School Quality and Student Outcomes: Evidence from Model Schools in India," 2020.

Lai, Fang, Renfu Luo, Linxiu Zhang, Xinzhe Huang, and Scott Rozelle, "Does Computer-Assisted Learning Improve Learning Outcomes? Evidence from a Randomized Experiment in Migrant Schools in Beijing," *Economics of Education Review*, 2015, *47*, 34–48.

Lee, David, "Training, Wages, and Sample Selection: Estimating Sharp bounds on Treatment Effects," *Review of Economic Studies*, 2009, *3* (76), 1071–1102.

Linden, Leigh, Abhijit Banerjee, and Esther Duflo, "Computer-Assisted Learning: Evidence from a Randomized Experiment," *Poverty Action Lab Paper*, 2003, *5.*

Mbiti, Isaac and Dominic Rodrigue-Segura, "Back to basics: Curriculum Reform and Student Learning in Tanzania," 2020.

_ , **Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani**, "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania," 2019, *24876.*

Muralidharan, Karthik, "Priorities for Primary Education Policy in India's 12th Five-Year Plan," *India Policy Forum*, 2013, *9* (1), 1–61.

_ , **Abhijeet Singh, and Alejandro J Ganimian**, "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India," *American Economic Review*, 2019, *109* (4), 1426–60.

_ **and Nishith Prakash**, "Cycling to School: Increasing Secondary School Enrollment for Girls in India," *American Economic Journal: Applied Economics*, 2017, *9* (3), 321–50.

— **and Venkatesh Sundararaman**, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 2011.

Murnane, Richard J. and Alejandro Ganimian, "Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Impact Evaluations," *Review of Educational Research*, 2016, *86* (3), 719–755.

Pratham, "Annual Status of Education Report," Technical Report, Pratham 2016.

Pritchett, Lant and Amanda Beatty, "The Negative Consequences of Overambitious Curricula in Developing Countries," *American Economic Journal: Applied Economics*, 2012.

Schell, Julie, Brian Lukoff, and Eric Mazur, "Catalyzing Learner Engagement Using Cutting-Edge Response Systems in Higher Education," 2013.

UNESCO, "More Than One-Half of Children and Adolescents Are Not Learning Worldwide," 2017.

West, Martin, Matthew Kraft, Amy Finn, and Angela Duckworth, "Promise and Paradox: Measuring Students Non-cognitive Skills and the Impact of Schooling," 2014.

World Bank, "World Development Report," 2018.

— , "World Bank Development Indicators," 2020.

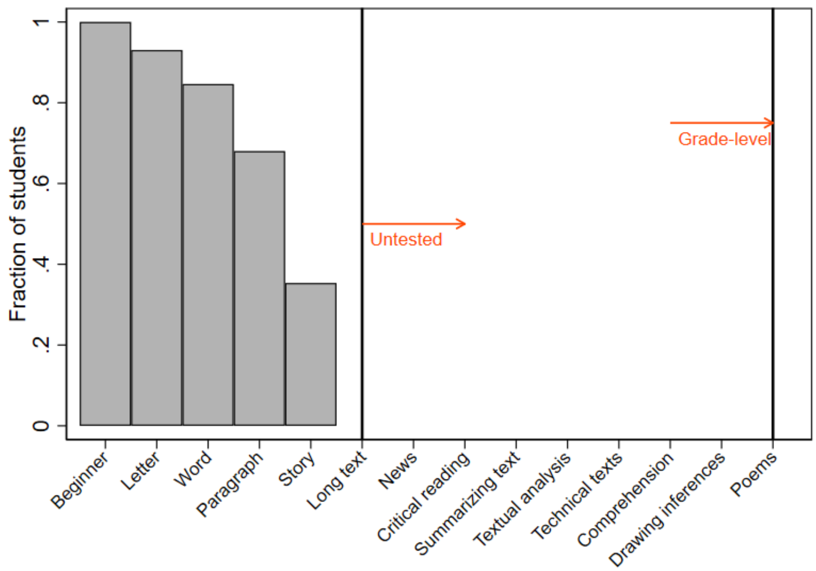
Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *The Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Zhang, Ping, Lin Ding, and Eric Mazur, "Peer Instruction in Introductory Physics: A Method to Bring about Positive Changes in Students Attitudes and Beliefs," *Physical Review Physics Education Research*, 2017.

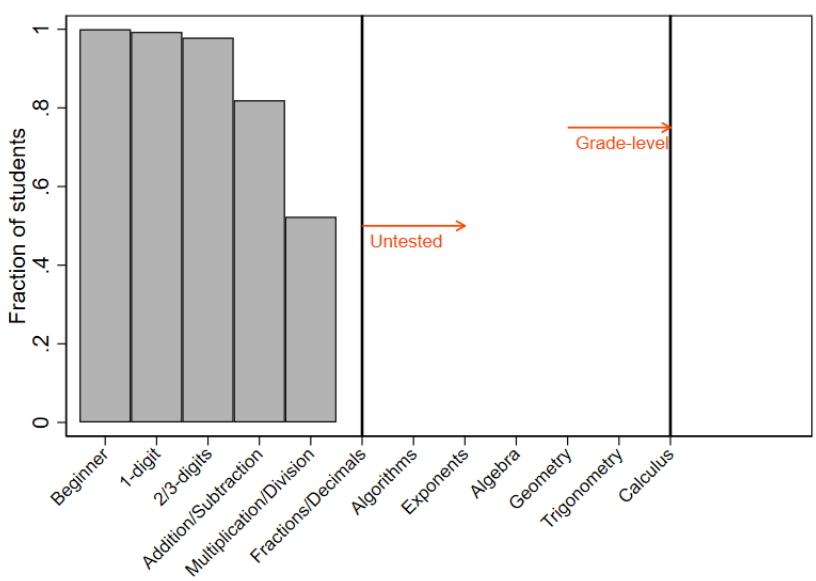
Figures and Tables

Figure 1: Reading and Math Levels at Baseline

(a) Reading scores



(b) Maths scores



Notes: The above figure shows the cumulative fraction of students across various levels of reading and math as measured by the ASER test.

Table 1: Sample Characteristics

	N	Whole Sample	Control	Treatment	(2)-(3)
		(1)	(2)	(3)	(4)
Panel A: Student characteristics					
Female	1044	0.412	0.409	0.415	-0.006
		[0.492]	[0.492]	[0.493]	
Age	1044	14.887	14.856	14.919	-0.063
		[0.867]	[0.892]	[0.841]	
Hindu	1044	0.761	0.741	0.781	-0.040
		[0.427]	[0.439]	[0.414]	
SC/ST	1044	0.402	0.396	0.409	-0.013
		[0.491]	[0.489]	[0.492]	
OBC	1044	0.451	0.455	0.448	0.007
		[0.498]	[0.498]	[0.498]	
Panel B: Guardian characteristics					
Illiterate	1044	0.200	0.203	0.198	0.005
		[0.400]	[0.402]	[0.399]	
Any schooling	1044	0.725	0.735	0.715	0.020
		[0.447]	[0.442]	[0.452]	
Income $<$ Rs. 10k	882	0.782	0.772	0.792	-0.020
		[0.413]	[0.420]	[0.406]	
Panel C: Student cognitive, non-cognitive skills and parental support					
ASER Reading	647	3.875	3.423	4.391	-0.968***
		[1.138]	[1.092]	[0.957]	
ASER Maths	647	4.349	4.174	4.550	-0.376***
		[0.824]	[0.914]	[0.654]	
Parent Support	810	0.100	0.198	-0.009	0.207**
		[1.300]	[1.348]	[1.237]	
Life Skills	810	0.015	0.017	0.013	0.004
		[1.789]	[1.732]	[1.853]	
Attitudes	810	0.021	0.063	-0.026	0.090
		[1.355]	[1.382]	[1.325]	

Notes: See Appendix B for measurement of ASER and Parent Support Index variables. SC/ST and OBC is an indicator variable that takes the value 1 if the student is from a Scheduled Caste/Tribe or Other Backward Caste respectively. Illiterate (Any schooling) takes the value 1 if the student's guardian is illiterate (has any schooling) and 0 otherwise. The omited category is if the student does not know his/her guardian's education. Income<10k takes the value 1 if the guardian's income is less than 10k and 0 otherwise. Column (4) reports the difference between treatment and control and the asteriks report a t-test of whether this difference is statistically significant. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 level.

Table 2: Impact on Cognitive Ability

	Standardized ASER Score		ASER Score	
	Reading	Maths	Reading	Maths
	(1)	(2)	(3)	(4)
	OLS	OLS	O. Logit	O. Logit
PANEL A: Without individual and school controls				
Treatment	0.528***	0.479***	1.299***	1.684***
	(0.133)	(0.102)	(0.323)	(0.368)
Baseline	0.581***	0.544***	1.274***	1.730***
	(0.0610)	(0.0476)	(0.157)	(0.199)
p-val (OLS)	0.00	0.00	0.00	0.00
p-val (RI)	0.00	0.00	0.00	0.00
R^2	0.471	0.465		
PANEL B: With individual and school controls				
Treatment	0.470***	0.516***	1.211***	1.887***
	(0.146)	(0.118)	(0.373)	(0.518)
Baseline	0.580***	0.533***	1.299***	1.727***
	(0.0606)	(0.0508)	(0.160)	(0.222)
p-val (OLS)	0.00	0.00	0.00	0.00
p-val (RI)	0.00	0.00	0.00	0.00
R^2	0.484	0.480		
N	647	647	647	647

Notes: The outcome variable in columns (1) and (2) are the ASER scores for Reading and Maths that have been standardized to have mean 0 and standard deviation 1 for the control group. Individual controls include gender, age, religion, and caste. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 respectively.

Table 3: Impact on Standardized National Exam Scores

	Tamil	English	Maths	Science	Social Science	Total
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Without Baseline ASER Controls						
Treatment	1.423	1.486	0.865	1.551	1.931	1.537
	(2.463)	(2.975)	(1.806)	(1.999)	(2.314)	(1.561)
p-val (OLS)	0.563	0.617	0.632	0.438	0.404	0.325
p-val (RI)	0.636	0.692	0.724	0.448	0.400	0.412
R^2	0.147	0.117	0.102	0.141	0.115	0.135
PANEL A: With Baseline ASER Controls						
Treatment	-1.392	-0.586	-0.239	0.122	-0.300	-0.396
	(2.244)	(3.041)	(2.015)	(1.921)	(2.361)	(1.579)
Baseline ASER	7.486***	5.511***	2.934***	3.801***	5.934***	5.140***
	(0.980)	(0.718)	(0.614)	(0.647)	(0.699)	(0.596)
p-val (OLS)	0.535	0.847	0.906	0.949	0.899	0.802
p-val (RI)	0.556	0.884	0.936	0.948	0.892	0.852
R^2	0.238	0.201	0.129	0.191	0.167	0.219
Control Mean	65.71	51.64	46.78	63.14	60.91	57.58
N	677	677	677	677	677	677

Notes: All outcome variables are the percentage marks scored by a student in their Class X exams. Both panels include individual and school controls. Individual controls include gender, age, religion, and caste. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Panel B controls for the baseline ASER score (Maths), while Panel A does not control for it. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 respectively.

Table 4: Impact on Students' Effort on Studies, Parental Support

	Absenteeism	Parent Support	Tuition		Self-Study	
			At least 1 day	Days/Week	At least 1 day	Days/Week
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Without individual and school controls						
Treatment	-0.0878	0.249	-0.00268	-0.0202	0.00438	-0.0264
	(0.180)	(0.312)	(0.0461)	(0.297)	(0.0441)	(0.136)
Baseline	0.172***	0.270***	0.544***	0.493***	0.250***	0.267***
	(0.0405)	(0.0331)	(0.0536)	(0.0474)	(0.0534)	(0.0645)
p-val (OLS)	0.626	0.424	0.954	0.946	0.921	0.846
p-val (RI)	0.636	0.500	0.952	0.944	0.904	0.856
R^2	0.025	0.064	0.260	0.232	0.056	0.044
PANEL B: With individual and school controls						
Treatment	-0.148	0.105	-0.0437	-0.300	-0.0115	-0.0561
	(0.225)	(0.348)	(0.0566)	(0.368)	(0.0600)	(0.136)
Baseline	0.154***	0.275***	0.536***	0.481***	0.239***	0.233***
	(0.0425)	(0.0335)	(0.0525)	(0.0480)	(0.0518)	(0.0703)
p-val (OLS)	0.511	0.762	0.440	0.415	0.847	0.680
p-val (RI)	0.460	0.780	0.388	0.332	0.840	0.756
R^2	0.047	0.089	0.279	0.253	0.073	0.077
Control mean	1.775	14.21	0.406	2.385	0.656	1.198
N	841	841	557	557	557	557

Notes: "Absenteeism" is the number of days in the last week that a student was absent from school. "Parent Support" is a composite index described in Appendix Section B. Tuition in Column (3) takes the value 1 if a student takes private tuition classes and is the number of days/week that the students attends tuition classes in Column (4). Self-study in Column (5) takes a value 1 if a student spent at least 1 day in the last week self-studying and 0 otherwise, and the number of days/week that a student self-studies in Column (6). Individual controls include gender, age, religion, and caste. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 respectively.

A Appendix Tables

Table A1: Impact by Baseline ASER Levels

	Reading level			Math Score		
Baseline level:	Word	Paragraph	Story	Digits	Subtraction	Division
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Without individual and school controls						
Treatment	0.379**	0.471***	0.290	0.477***	0.661***	0.324***
	(0.154)	(0.166)	(0.182)	(0.165)	(0.106)	(0.120)
p-val (OLS)	0.014	0.005	0.11	0.004	0.00	0.007
p-val (RI)	0.008	0.00	0.22	0.00	0.00	0.000
R^2	0.099	0.101	0.022	0.092	0.240	0.095
PANEL B: With individual and school controls						
Treatment	0.247	0.224	0.281	0.499**	0.718***	0.326**
	(0.221)	(0.194)	(0.191)	(0.219)	(0.146)	(0.127)
p-val (OLS)	0.263	0.247	0.141	0.02	0.00	0.010
p-val (RI)	0.084	0.084	0.348	0.00	0.00	0.012
R^2	0.228	0.206	0.098	0.221	0.275	0.131
N	194	221	232	100	201	346

Notes: Each column restricts the sample of students to that level of reading/maths at baseline and examines the impact of the treatment on endline scores. Individual controls include gender, age, religion, and caste. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 respectively.

Table A2: Impact on Parental Support

	Total	Events	Homework	School	Talk
	(1)	(2)	(3)	(4)	(5)
PANEL A: Without individual and school controls					
Treatment	0.249	0.0685	0.0157	0.0509	0.0661
	(0.312)	(0.105)	(0.106)	(0.155)	(0.123)
Baseline	0.270***	0.233***	0.178***	0.0726***	0.213***
	(0.0331)	(0.0324)	(0.0319)	(0.0266)	(0.0329)
p-val (OLS)	0.424	0.516	0.882	0.743	0.592
p-val (RI)	0.500	0.560	0.904	0.748	0.580
R^2	0.064	0.049	0.028	0.005	0.041
PANEL B: With individual and school controls					
Treatment	0.105	0.0176	0.0526	0.0116	0.00535
	(0.348)	(0.112)	(0.141)	(0.188)	(0.172)
Baseline	0.275***	0.239***	0.176***	0.0693***	0.214***
	(0.0335)	(0.0336)	(0.0316)	(0.0251)	(0.0342)
p-val (OLS)	0.762	0.875	0.709	0.951	0.975
p-val (RI)	0.780	0.864	0.744	0.952	0.984
R^2	0.089	0.064	0.038	0.027	0.052
N	841	841	841	841	841
Control mean	14.21	3.560	3.998	3.328	3.328

Notes: Each outcome variable is a measure from 1 (Never)–5 (Always) on how often a parent supervises the studies of the child. Individual controls include gender, age, religion, and caste. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes p<0.1, ** denotes p<0.05 and *** denotes p<0.01 respectively.

B Program Description, ASER Measurement

B.1 Components of Avanti's Program

Table B1: Avanti Program Components Description

Technology: 15 short five-minute videos in the local language that explain concepts through the use of real world examples.	1 hour 15 minutes (15 videos) per week
Peer-learning: Students organized in groups of 6 and tasked with collaboratively solving worksheets. Facilitators intervene to encourage discussion and resolve questions only when necessary.	1 hours 30 minutes (6 worksheets) per week
Trained Facilitators: Facilitators (from the local area) led lectures to summarize key concepts based on facts emerging from an in-class discussion.	1 hour 40 minutes (20 min per day) per week

B.2 ASER Reading and Maths Survey

Table B2: Curriculum Mapping for ASER Survey

ASER Levels	**ASER Reading**	**Grade level**	**ASER Maths**	**Grade level**
Level 1	Beginner		Beginner	
Level 2	Identifying letters	< Grade 1	Number recognition 0 - 9	< Grade 1
Level 3	Reading words	< Grade 1	Number recognition 10 - 99	< Grade 1
Level 4	Reading a paragraph	Grade 1	Subtraction with borrowing	Grades 2 or 3
Level 5	Reading a story	Grade 2	Division with remainder	Grades 3 or 4

B.3 Parental Support Index

Table B3: Parent Support Questions

Variable	**Survey Question**
School	How often do your parents help you with your school work?
Talk	How often do your parents talk to you about what you are doing in school?
Homework	How often do your parents ask you about homework?
Events	How often do your parents go to meetings or events at school?

C Attrition

Of the 1044 students in our baseline sample we only have 841 students in a the endline (an attrition of 20.3 percent). We investigate the differential attrition across treatment and control schools on observable characteristics as well as its impact on our treatment estimates in two ways. First, we report whether there is differential attrition between treatment and control. We estimate the following specification:

$$M_i = \alpha + \beta X_i + \gamma T_i + \delta T_i \times X_i + \varepsilon_i$$

where M_i is a dummy variable that takes the value 1 if the student stays in the sample and 0 if the student drops out between the baseline and endline. δ is the coefficient of interest since it tells us whether characteristics of students who stay in the sample is statistically different on average between control and treatment schools. We report the estimates for δ in Table C1. As reported in the table, attrition is small in magnitude and statistically insignificant between treatment and control schools for almost all student characteristics (except age and religion, which we use as controls in our analysis).

Table C1: Differential Attrition between Control and Treatment

	Coefficient	S.E.	p-value
	(1)	(2)	(3)
Female	0.04	0.06	0.49
Age	0.08	0.03	0.02
Hindu	0.17	0.07	0.02
SC/ST	0.00	0.10	0.99
OBC	-0.01	0.10	0.92
Illiterate	-0.02	0.08	0.80
Income $<$ Rs. 10k	-0.09	0.08	0.28
ASER Reading	0.00	0.02	0.97
ASER Maths	-0.03	0.02	0.25
Life Skills	-0.01	0.01	0.31
Student Attitudes	0.00	0.03	0.97
Parental Support	-0.04	0.04	0.37

Similarly, we also investigate the differential propensity between students in the control and treatment schools to take the ASER survey at baseline. As reported in Table C2, the differences are both statistically insignificant and small in magnitude.

Table C2: Differential Attrition in taking ASER Survey

	Coefficient	S.E.	p-value
	(1)	(2)	(3)
Female	0.02	0.05	0.67
Age	-0.03	0.03	0.34
Hindu	-0.09	0.06	0.09
SC/ST	0.02	0.08	0.84
OBC	-0.03	0.08	0.68
Illiterate	0.00	0.06	0.99
Income < Rs. 10k	0.06	0.06	0.39
Life Skills	0.01	0.01	0.62
Student Attitudes	0.01	0.02	0.63
Parental Support	0.00	0.01	0.80

In addition, to account for potential endogenous attrition from the sample, we also estimate Lee bounds on the treatment effects on the ASER test scores and the Grade 10 exam scores with bootstrapped standard errors (Lee, 2009). We use the baseline values as well as individual and school controls to tighten the bounds. As reported in Table C3, the lower bound for ASER learning is still around 0.16-0.17σ for Reading and Maths (though statistically insignificant at conventional levels), while the bounds for the Grade 10 exams are very small in magnitude (± 1 percentage point).

Table C3: Lee Bounds on Treatment Effects

	Lower bound	S.E.	Upper bound	S.E.
	(1)	(2)	(3)	(4)
ASER Reading	0.173	(0.183)	0.471***	(0.133)
ASER Maths	0.164	(0.242)	0.393***	(0.104)
Tamil	0.223	(3.414)	-1.348	(2.694)
English	0.0663	(2.561)	-0.577	(3.633)
Maths	1.118	(2.457)	-0.669	(3.170)
Science	0.808	(2.448)	-0.543	(2.155)
SSC	1.385	(3.125)	-0.724	(2.872)
Total	0.646	(2.217)	-0.725	(2.442)