Play to Learn - Improving Foundational Learning with Technology Aided Formative Teaching

Ramiro Burga, Divya Pandey, Sheetal Sekhri, Gauri Kartini Shastry*

December 20, 2022

Abstract

Computer-aided learning (CAL) can play a vital role in reducing learning gaps in developing countries by providing low-cost access to supplemental learning. While existing studies have focused on the effects of computer-based learning systems on older students, the impact of CAL on younger children during their formative years is still unknown. In this study, we will evaluate the effect of a formative teaching CAL model for foundational learning where teachers in grades 1 and 2 will leverage a game-based software, Chimple, to assign supplemental at-home learning activities.

Keywords: Education, Interactive app-based learning, Computer-aided learning, Primary school, Test-scores, Elementary school, Productivity in education, Efficiency in teaching, Educational software

JEL Classification: C93, I21, I24, J24, O15

*Affiliations: Burga, Acumen LLC; Pandey, International Initiative for Impact Evaluation (3ie), Delhi Office; Sekhri, University of Virginia; Shastry, Wellesley College. We are immensely grateful to the Central Square Foundation (CSF), the Chimple team, Bharti Foundation and the Awadh Research Foundation (ARF).

1 Introduction

The role of educational technology (ed-tech) or computer-aided learning (CAL) in bridging educational gaps has garnered substantial interest from researchers and policymakers (see Escueta et al. 2017 for a review). CAL has the potential to alter learning outcomes by expanding access to education, augmenting traditional classroom training, delivering personalized materials, and encouraging engagement through compelling content. Some studies have shown that ed-tech improves students' educational achievements significantly by providing adapting learning opportunities (Angrist, Bergman, and Matsheng 2022; Muralidharan, Singh, and Ganimian 2019). Nevertheless, the efficacy of such educational models has not been tested on younger children during their most formative years. In addition, the delivery model has either been phone and SMS or an app used in centers. Teacher-assisted remedial use of ed-tech where formative assessment is central has not been used.

In this study, we will examine the effect of a teacher-assisted ed-tech formative learning model on first and second grade students' test-scores in India, by leveraging the deployment of a free educational app named Chimple. The app is available in local vernacular (Hindi), and focuses on foundational learning (literacy and numeracy) with a game-based approach, which can boost children's engagement. Additionally, the Chimple dashboard tracks students' engagement levels and progress which aids formative assessment rather than diagnostic assessment. Given that the app can be easily installed on low-end smartphones, and is thereby accessible to low-income populations, our study will shed light on the effectiveness of CAL in improving the educational outcomes of children who come from marginalized backgrounds.

To assess the impact of Chimple on children's test-scores, we use a stratified randomized control trial conducted, where a grade is the unit of randomization. Our sample consists of students in grades 1 and 2 in 34 schools managed by the Bharti Foundation, across 5 districts in Haryana, India. We used baseline average test scores to group schools in 8 strata. In each stratum, 50% of schools were randomly chosen to offer the Chimple program in grade 1 (e.g., Schools A and B); the other schools offered the Chimple program in grade 2 (e.g., Schools C and D). Grade 1 in Schools A and B will serve as a comparison for Grade 1 in Schools C and D, and similarly, Grade 2 in Schools C and D will serve as a comparison for Grade 2 in Schools A and B.

The teachers of the treatment group will integrate Chimple into their course plans as a supplemental tool. In the treatment group, each week teachers assign homework via Chimple and track students progress via the dashboard. As schools have returned to in-person learning after the Covid-19 pandemic, our study will highlight how technology-aided supplemental learning can play an instrumental role in helping teachers address a heterogeneous classroom.

In addition to CAL models, sharing information regarding children's school activities with parents at a high-frequency could alter educational outcomes (Barrera-Osorio et al. 2020). Therefore, our study embeds another experiment to evaluate a teacher-driven model where students are encouraged to continue using the software because their teacher assigns content regularly and send reminders/nudges to the parents' mobile phones. To study how engagement responds to reminders, we randomly assigned parents into two groups, with one group receiving nudges infrequently (twice per month) and the other receiving nudges at a higher frequency (twice per week).

2 Intervention and Research Design

2.1 Chimple Learning Application

Our study leverages the deployment of a free educational app named Chimple, available in local vernacular language (Hindi in our case) geared towards foundational learning (literacy and numeracy) harnessing game-based approaches. Chimple can be installed on low-end smartphones and works offline with its library of 70 unique games - 50 for literacy and 20 for numeracy. The literacy ladder takes the child from letter recognition to sentence formation and comprehension. The Numeracy ladder takes the child from number sense to basic multiplication. Chimple has a dashboard that tracks the child's participation and progress. Chimple corresponds to the syllabi for the grade proposed by the government (K-2). These features make it very attractive to marginalized and vulnerable populations as well. The game-based approach circumvents the issue of short attention spans of young children.

XPRIZE and UNESCO have tested Chimple in a pilot evaluation in Tanzania. This program lasted 15 months (December 2017 -March 2019) with 500 children aged 7-11 across 30 villages in the Tanga region of Tanzania. The product was accessed via tablets. Learning outcomes were assessed via EGRA/EGMA test (RTI) and Social-Emotional assessment (UNESCO) as well. Large effects of over 0.5 standard deviations were observed for both literacy and numeracy.

Chimple in India: Implementation in India was initiated in a computer lab setting in 15 schools in Mumbai, Maharashtra, in partnership with the Brihan Mumbai Municipal Corporation, although this was disrupted due to COVID. Around 500 children from low income households accessed Chimple via Door Step School Mumbai community centers. But the efficacy of Chimple in a teacher-driven at-home environment has not been tested.

2.2 Setting

The study will be implemented in 34 schools operated by Bharti Foundation. Bharti Foundation runs 38 schools in the state of Haryana, 4 of which were excluded from the study since they were used to deploy a pilot version of Chimple.[1] Total enrollment across all 38 schools is approximately 1000 students per grade. More than 80 percent of children enrolled are of disadvantaged castes, and 50 percent of households earn less than USD 1,300 per year. Approximately 90 percent of households have access to at least one smartphone that can run Chimple. Figure 1 plots the spatial distribution of the 34 schools participating in this study.

2.3 Study Design and Interventions

Our study encompasses two randomized experiments. In the first randomization, we assess the impact of Chimple on children's test scores using a stratified randomized control trial, where a class is the unit of randomization. The sample includes 34 schools with 1 class each in grades 1 and 2, Chimples targeted age group. The grade randomization was done at the end of March 2022–before the start of the 2022-2023 school year– based on school administrative records from the previous academic year. Specifically, we used average test scores from the end of 2021-2022 school year to group schools in 8 strata (7 groups of 4 schools and one group of six schools). In each stratum, 50% of schools were randomly chosen to offer the Chimple program in grade 1 (e.g., Schools A and B); the other schools will offer the Chimple program in grade 2 (e.g., Schools C and D). Grade 1 in Schools A and B will serve as a comparison for Grade 1 in Schools C and D, and similarly, Grade 2 in Schools C and D will serve as a comparison for Grade 2 in Schools A and B. Table 1

1. These four schools were selected based on their proximity to Delhi. We used these schools to pilot all surveys and assessments of this study.

shows that schools whose treated grades are first and second grade are not different in observable characteristics measured at the end of the 2021-2022 school year.

Teachers in treated grades will integrate Chimple into their course plans as a supplemental tool. Each week teachers will assign homework via Chimple and track students progress via the dashboard. Ninety percent of students have access to Chimple via a parent's phone, but teachers are also expected to help students who have limited access to Chimple at home. The intervention will leverage Whatsapp, a messaging tool, as a platform for teachers to ask parents to install Chimple on their phones. Schools will run a parent information session to highlight the potential value of Chimple for their childrens learning.

For our second experiment, we randomized the intensity of text-based nudges across students from treated classrooms. Using the enrollment list from the end of May, students were randomly assigned to a low or high-intensity WhatsApp group in mid-June. The randomization was stratified by classroom and parental access to smartphones (access or no access). In other words, we randomized the WhatsApp group for each classroom-by-smartphone access cell. Table 2 indicates that students in the high- and low-intensity groups share similar observable characteristics.

Households in the high-intensity group receive two text messages per week while those in the low-intensity group receive a text message twice a month.[2] The content of the text messages is focused on reminding parents that their children should complete the tasks posted by teachers on the Chimple app, while at the same time highlighting the potential benefits of Chimple. An example of text messages is:[3]

> "Hello! Your childs teacher has posted weekly learning activities on Chimple. Chimple uses games and simple language to teach children basic concepts in an engaging way. Please encourage your child to complete them in a timely manner."

2.4 Power calculation

In this section, we report the Minimum Detectable Effects (MDEs) corresponding to the number of grades and number of students included in this study. Given the relatively small samples of our

2. Messages are sent on Tuesday and Friday of every week for the high-intensity group and Tuesdays every two weeks for the low-intensity group. The first message to the low and high-intensity groups was sent on September 9th and the messages will continue until the end of March.

3. Other variations can be found in Appendix A.

data, we use the randomized inference approach (Splawa-Neyman, Dabrowska, and Speed 1990, Athey and Imbens 2017).[4]

Using data from past learning assessments conducted by Bharti Foundation, we estimate that the minimal detectable effect (MDE) in the classroom-level randomization ranges from 0.2 to 0.3 SD. That is, we have a reasonable chance (80% power) of statistically detecting a significant difference (5% significance) in test scores between children in treatment classes and children in comparison classes if the true effect of Chimple on test scores is 0.2 - 0.3 SD. An effect of 0.3 SD would be considered a large effect in the education literature; at the same time, the pilot evaluation of Chimple found impacts of approximately 0.5 SD.

We used past learning assessments and dashboard data on weekly Chimple engagement (in minutes) to estimate the MDE in the student-level randomization of WhatsApp nudges. It is essential to mention that we used usage information from weeks before we started to nudge families.[5] We obtained MDEs in the ranges of 0.14-0.15 SD and 0.19-0.2 SD for test scores and weekly engagement, respectively.

3 Data

3.1 Student Test-scores

This study's primary outcome of interest is students' test scores on English and math tests administered by an independent organization (Awadh Research Foundation, ARF) and designed by the research team. The endline tests will be conducted in Bharti schools in early March 2023. To minimize attrition, ARF will follow up with students who were not present in school on the day of the test. The research team will design the test questions following recommendations in the Early Grade Mathematics Assessment (EGMA) toolkit from RTI International and include questions similar to those in tests conducted by Pratham-ASER. The test will follow the same format and difficulty as tests administered by Bharti.

Due to various constraints, the research team was not able to conduct such tests before treatment assignment; however, we were able to conduct tests following this protocol in September 2022. The

4. The formulas we used can be found in Appendix B.

5. We used the number of minutes students engaged in the app during the August 7th-August 13th week. Results are almost identical if we use any other week before we begin nudging parents.

school year begins in April-May but school closes for extended periods in June and enrollment does not stabilize until July-August. Our baseline test was completed after Bharti's internal assessments were completed and before we began the parent nudges. We designed the baseline test as we anticipate designing the endline test, described above.

We will also collect test scores measured by Bharti, internally. Bharti collects two assessments every school year for all elementary-school grades. The first one takes place in the middle of the academic year, and the second at the end. We will use Bharti's test scores from the end of the 2022-2023 school year assessment as an alternative outcome variable. In addition, we can use Bharti test scores from the 2021-2022 school year as baseline scores but only for students who were enrolled with Bharti in the previous year, either first or second grade.

3.2 Chimple Data System

An important intermediate outcome for the nudging intervention is take-up from both teachers (in assigning Chimple activities) and students (in engaging with Chimple). The Chimple system logs all student and teacher interactions with the App. Chimple provided the data to the research team through a dashboard, from which we downloaded the data in excel files for a pre-specified time window. On the student side, the dashboard data includes the time spent per week and the number of activities completed by students for each subject (English, Hindi, and Math). Similarly, the teacher data on the dashboard includes the time spent and the number of activities assigned by teachers for each subject.

3.3 School Administrative Records

We have access to administrative records for the 2020-2022 and 2022-2023 school years. The information contained in these records includes student test scores and demographics, data on at-home device access from a survey conducted by Bharti, and teacher characteristics.

3.4 Teacher survey

We conducted zoom interviews in November-December 2022 with grade 1 and 2 teachers from all schools participating in the study. This survey inquires teachers about teaching practices including homework assignment, interaction with parents, Chimple usage and perceptions.

3.5 Household Endline Survey

The endline survey sample has been drawn from the student enrollment list from May 2022. We randomly selected 10 households per grade to be interviewed. Since one ex-ante goal was to look at heterogeneity by device access, the selection stratified students by classroom and phone access. In the endline survey, we will ask parents about their perceptions of Chimple as a learning tool. We will also gauge children's and parent's familiarity with Chimple to determine whether parents were helping children with their Chimple homework. We will also inquiry students and parents from control shools about Chimple usage to identify potential spillovers.

4 Regression Specification

4.1 Intention-to-treat estimates

To estimate the impact of the Chimple app on students' performance , we will estimate intention-to-treat (ITT) effects using the following regression specification:

$$Y_{igst_1} = \beta_0 + \beta_1 Treat_{gs} + \mu_k + \lambda_r + Y_{igst_0} + X^{'}_{gst_0}\Phi + \epsilon_{igst_1} \tag{1}$$

where Y_{igst_1} is the test-score of student i in grade g in school s at the time of endline t_1. Test scores will be normalized using the distribution of the Control group. Here, $Treat_{gs}$ indicates that grade g of school s was selected to offer the Chimple program, and β_1 is the ITT estimate. Our regressions will control for strata and district fixed effects (μ_k, λ_r). We will also control for students baseline test-scores (Y_{igst_0}) and other predetermined characteristics (X_{gst_0}).[6] ϵ_{igst_1} is the error term clustered at the school-grade level. We will report estimates on students' performance in Mathematics and English tests separately.

Note that β_1 in regression 1 offers the average ITT impact of the Chimple app across children in the high-intensity group and the low-intensity group. To separate these effects, we will estimate the following regression specification:

6. Recall that we have two possible baseline controls: 1) we have data from Bharti from the previous school-year for a subset of the children and 2) we have data from our own test administered a few months after treatment assignment. We will control for normalized measures of these scores separately, including a control to indicate missing observations (which will be inputted 0).

$$Y_{igst_1} = \beta_0 + \beta_1 H_{igs} + \beta_2 L_{igs} + \mu_k + \lambda_r + Y_{igst_0} + X'_{gst_0}\Phi + \epsilon_{igst_1} \tag{2}$$

where H_{igs} and L_{igs} are indicators for whether a student in the treatment group belongs to the high-intensity and low-intensity group.

4.2 Impact of nudges

We will explore the impact of high-intensity nudging on measures of take-up from the Chimple dashboard with the following regression specification estimated on students in Chimple classes:

$$M_{igs} = \beta_0 + \beta_1 H_{igs} + \mu_k + \lambda_r + X'_{gst_0}\Phi + \epsilon_{igs} \tag{3}$$

where M_{igs} is the average number of minutes a student spent on the app per week and other variables are as defined above. We will also exploit the high-frequency nature of the data and the timing of the nudges in the following regression specification at the student-day level:

$$M_{igsd} = \beta_0 + \beta_1 N_{igsd_{-1}} + \beta_2 N_{igsd_0} + \beta_3 N_{igsd_{+1}} + \mu_k + \lambda_r + \nu_d + X'_{gst_0}\Phi + \epsilon_{igsd} \tag{4}$$

where M_{igsd} is the number of minutes a student spent on the app on day d and our primary independent variables indicate whether the household received a nudge on that day (N_{igsd0}), the day before (N_{igsd-1}) or the day after (N_{igsd+1}). ν_d are day fixed-effects. Recall that the low-intensity group also receives nudges, twice a month. Thus, to estimate the impact of additional nudges, we will estimate separate versions of this regression either i) interacting these nudge indicators with indicators for being in the high intensity group or ii) interacting these nudge indicators with the cumulative number of previous nudges. Early on in the intervention, β_1 may be expected to be 0 but as the household receive more nudges, they may anticipate the timing of the nudges.

We will estimate regression equations 3 and 4 using the number of activities and assignments completed as the dependent variable as well, although with the caveat that the number of assignments completed depends on how many assignments are assigned by the teacher.

4.3 Local average treatment effects

Using these measures of take-up, we can estimate local average treatment effects (LATE) of Chimple usage on test scores with the following regression specification:

$$Y_{igst_1} = \beta_0 + \beta_1 M_{igs} + \mu_k + \lambda_d + Y_{igst_0} + X'_{gst_0}\Phi + \epsilon_{igst_1} \tag{5}$$

where H_{igs} and L_{igs} will be used as instruments for M_{igs}. We will estimate this in two ways: first, we will focus on students in the Chimple group, using variation from the impact of the nudges. Second, if we can verify in the endline survey that control households were not using Chimple, we will define M_{igs} as 0 for the control group and use the entire sample. In an alternate specification, we will use the fraction of weeks a child engaged with the Chimple app as a measure of take-up. β_1 captures the LATE estimate.

4.4 Heterogeneity

4.4.1 Heterogeneity by Student Characteristics

We will investigate whether our ITT effects vary by students' gender, socioeconomic status, at-home device access and baseline test-scores. Since the baseline data on test-scores was collected after the start of our intervention to offer Chimple and before the nudges intervention, we will also explore heterogeneity by pre-treatment assignment test-scores where possible. We will use indicators for whether students belong to lower castes (OBC/SC) and whether they belong to high, middle or lower income households to examine heterogeneity by socioeconomic status.

4.4.2 Heterogeneity by Teacher Characteristics

Existing literature has shown that teachers' characteristics play a crucial role in determining students' achievement (Bettinger and Long 2005; Dee 2004). In our study, we will explore how the interaction between teachers' characteristics and our treatment drive students' performance. More specifically, we will explore heterogeneity in our estimates by teachers' sex, education level, years of experience and age.

4.4.3 Heterogeneity by Intensity of Covid

We will attempt to ascertain whether the intensity of the Covid-19 pandemic exacerbated the impact of the Chimple app. We aim to identify sources of shocks by collecting data on the Covid-induced death rates in the school districts or deaths in the student's immediate family. We will examine heterogeneity by Covid impacts on student's test-score, by interacting equations (1) and (2) with indicators for Covid-specific deaths in the student's family. Another way to identify Covid intensity is through collecting district-level data on Covid-related deaths and classifying districts into above median Covid shocks and below median Covid shocks based on the death rates in all districts of the state of Haryana. We could also identify the intensity of the Covid shock using administrative data on whether a district was identified as red, yellow or green during Covid and the duration of the lockdown.[7] Our ability to conduct this analysis will depend on the availability of data on Covid-related deaths at the household and/or district-level.

Finally, we do not rule out the possibility of exploring impacts on other measures for student's engagement on the Chimple app and other dimensions of heterogeneity. In the final paper, we will make a clear distinction between tests that were included in the PAP and those which were not.

4.4.4 Spillovers

Bharti Foundation has asked the teachers across the grades not to discuss the intervention. We are also going to collect data from teachers and students about a discussion of Chimple with others ask student parents at end-line if the student used Chimple for both the treatment and control group to determine and measure spillovers if any.

5 Analysis Prior to Drafting PAP

We drafted this PAP after Chimple was deployed in the treatment schools and the baseline survey had been conducted. We had also started sending the high and low-intensity nudges before drafting this PAP. To monitor the take-up of the Chimple app in a given week we looked at the following descriptive statistics:

7. A district was classified as red, orange or green based on whether there were many, few or zero Covid cases in it.

(a) The average time spent on the app across all schools.

(b) The average number of assignments given on the app across all schools.

(c) The average number of assignments completed on the app across all schools.

Nevertheless, we did not explore any differences in the above described variables by our interventions.

References

Angrist, N., P. Bergman, and M Matsheng. 2022. “Experimental evidence on learning using low-tech when school is out.” *Nature Human Behavior* 6:941950.

Athey, Susan, and Guido W Imbens. 2017. “The econometrics of randomized experiments.” In *Handbook of economic field experiments,* 1:73–140. Elsevier.

Barrera-Osorio, Felipe, Kathryn Gonzalez, Francisco Lagos, and David J. Deming. 2020. “Providing performance information in education: An experimental evaluation in Colombia.” *Journal of Public Economics* 186:104185. ISSN: 0047-2727. https://doi.org/https://doi.org/10.1016/j.jpubeco.2020.104185. https://www.sciencedirect.com/science/article/pii/S0047272720300499.

Bettinger, Eric P., and Bridget Terry Long. 2005. “Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students.” *American Economic Review* 95 (2): 152–157. https://doi.org/10.1257/000282805774670149. https://www.aeaweb.org/articles?id=10.1257/000282805774670149.

Dee, Thomas S. 2004. “Teachers, Race, and Student Achievement in a Randomized Experiment.” *The Review of Economics and Statistics* 86 (1): 195–210. https://ideas.repec.org/a/tpr/restat/v86y2004i1p195-210.html.

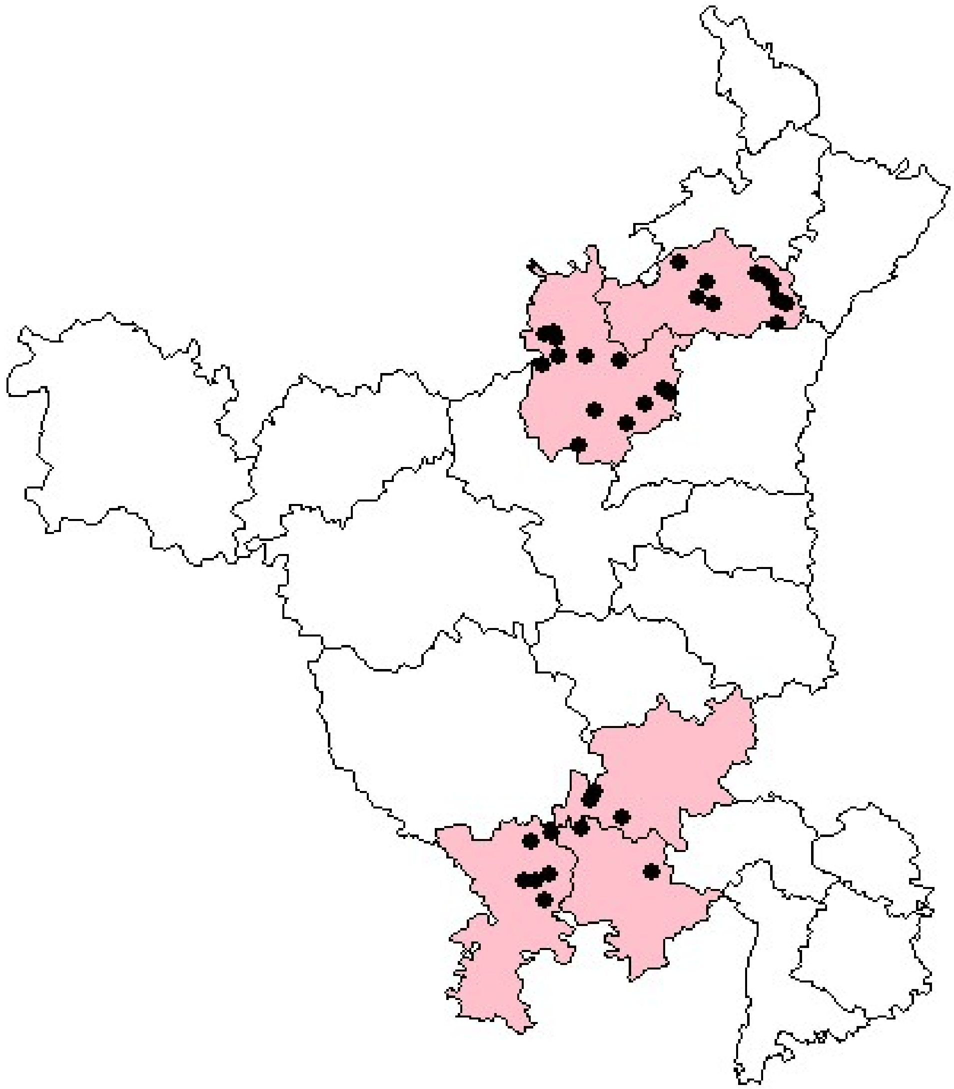
Escueta, Maya, Vincent Quan, Andre Joshua Nickow, and Philip Oreopoulos. 2017. *Education Technology: An Evidence-Based Review.* NBER Working Papers 23744. National Bureau of Economic Research, Inc, August. https://ideas.repec.org/p/nbr/nberwo/23744.html.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J. Ganimian. 2019. “Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India.” *American Economic Review* 109 (4): 1426–60. https://doi.org/10.1257/aer.20171112. https://www.aeaweb.org/articles?id=10.1257/aer.20171112.

Splawa-Neyman, Jerzy, Dorota M Dabrowska, and TP Speed. 1990. “On the application of probability theory to agricultural experiments. Essay on principles. Section 9.” *Statistical Science,* 465–472.

6 Figures

Figure 1. Spatial Distribution of Bharti Schools in Haryana



Note: The polygons in the map correspond to Haryana districts. Shaded districts contain Bharti schools. The dots indicate the location of Bharti schools.

7 Tables

Table 1. Balance on observable characteristics

	Grade 1 (1)	Grade 2 (2)	Standard P-value (3)	RI P-value (4)
Total enrollment	148.529	154.118	0.563	0.553
% male teacher	0.248	0.321	0.156	0.163
% general teacher	0.508	0.522	0.904	0.906
% OBC teacher	0.404	0.331	0.445	0.474
% SCST teacher	0.087	0.147	0.242	0.261
% MA teacher	0.344	0.471	0.151	0.146
% BA teacher	0.211	0.242	0.576	0.569
Average teacher age	36.229	37.560	0.333	0.319
% access chimple	0.910	0.909	0.975	0.972
HH Phone hours	2.395	2.582	0.659	0.683
% male students (grade 1)	0.513	0.535	0.472	0.472
% general caste students (grade 1)	0.179	0.167	0.864	0.847
% OBC caste students (grade 1)	0.364	0.364	0.996	0.996
% SC caste students (grade 1)	0.457	0.469	0.871	0.874
Avg. birth years (grade 1)	2015.391	2015.374	0.685	0.690
% migrant students (grade 1)	0.027	0.026	0.891	0.897
% High income students (grade 1)	0.617	0.602	0.907	0.903
% below PL students (grade 1)	0.185	0.174	0.858	0.859
mean score in SDs (grade 1)	-0.173	-0.012	0.135	0.150
% male students (grade 2)	0.467	0.541	0.100	0.092
% general caste students (grade 2)	0.174	0.172	0.973	0.968
% OBC caste students (grade 2)	0.339	0.354	0.799	0.780
% SC caste students (grade 2)	0.487	0.474	0.845	0.848
Avg. birth years (grade 2)	2014.366	2014.369	0.937	0.939
% migrant students (grade 2)	0.047	0.037	0.663	0.669
% High income students (grade 2)	0.702	0.657	0.696	0.694
% below PL students (grade 2)	0.224	0.186	0.477	0.467
mean score in SDs (grade 2)	-0.068	-0.174	0.286	0.282
mean score in SDs in 2021 (grade 1)	0.060	-0.067	0.511	0.497
mean score in SDs in 2021 (grade 2)	0.102	-0.051	0.539	0.5630
Number of schools	17	17	34	34

Note: In this table, the unit of observation is the school. RI stands for Randomized inference. P-values in column 3 were calculated using strata fixed effects. The same strata were used to calculate p-values with RI. We used the *ritest* command in the Stata software for RI calculations.

Table 2. Balance on observable characteristics-student randomization

	High-Frequency		Low-Frequency		Standard	RI
	N	mean	N	mean	P-Value	P-value
	(1)	(2)	(3)	(4)	(5)	(6)
Male	406	0.55	406	0.50	0.25	0.29
Age	406	6.11	406	6.14	0.18	0.16
Migrant Family	406	0.02	406	0.03	0.99	1.00
Father occupation: Agriculture	405	0.10	405	0.09	0.68	0.74
Father occupation: Labor	405	0.75	405	0.74	0.96	0.96
Father occupation: Private Job	405	0.07	405	0.07	0.96	1.00
Family income: Below 50 000	405	0.34	406	0.35	0.58	0.50
Family income: 50 000 to 100 000	405	0.52	406	0.50	0.51	0.53
Family income: Above 100 000	405	0.14	406	0.14	1.00	1.00
Mother's occupation: Wife	404	0.97	405	0.97	0.66	0.59
BLP family	406	0.16	406	0.17	0.68	0.68
Cast: Gen	406	0.17	406	0.17	0.85	0.93
Cast: OBC	406	0.38	406	0.33	0.04**	0.14
Cast: SC	406	0.45	406	0.50	0.19	0.17

Note: In this table, the unit of observation is the student. RI stands for Randomized inference. P-values in column 5 were calculated using strata fixed effects. The same strata were used to calculate p-values with RI. We used the *ritest* command in the Stata software for RI calculations.

Appendixes [Not for publication]

A Nudges: WhatsApp Messages

- Hello! Your childs teacher has posted weekly learning activities on Chimple. Chimple uses games and simple language to teach children basic concepts in an engaging way. Please encourage your child to complete them in a timely manner.
- Hello! Your childs teacher has posted weekly learning activities on Chimple. Chimple uses games and simple language to teach children basic concepts in an engaging way. If you encourage your child to complete these activities, it can improve his or her grasp of these basic concepts.
- Hello! Your childs teacher has posted weekly learning activities on Chimple. Chimple uses games and simple language to teach children basic concepts in an engaging way. Continued usage can be very important for learning. So please do not remove the app after usage.

B Power Calculation Formula

$$MDE = [\Phi^{-1}(\beta) + \Phi^{-1}(1-\alpha/2)] \times [\hat{V}_{strat}(\hat{\tau})]^{0.5} \tag{6}$$

$$\hat{V}_{strat}(\hat{\tau}) = \sum_{g=1}^{G} \hat{V}(\hat{\tau}_g) \times \left(\frac{N_g}{N}\right)^2 \tag{7}$$

$$\hat{V}(\hat{\tau}_g) = \left(\frac{s_{t,g}^2}{N_{t,g}}\right) + \left(\frac{s_{c,g}^2}{N_{c,g}}\right) \tag{8}$$

Where:

- β: Power of the test.
- α: Significance level.
- $\hat{V}(\hat{\tau}_g)$: Within stratum-variance for stratum g.
- $s_{(t,c),g}^2$: Variance of treated/control units in stratum g.
- $N_{(t,c),g}$: Number of treated/control units in stratum g.
- N_g: Number of units in stratum g.
- N: Total number of units.