Steven J. Green
School of International
& Public Affairs
Department of Economics

# Boosting Study Habits with High-Frequency Information: A Field Experiment to Aid Disadvantaged Students

Tomoki Fujii

Christine Ho

Rohan Ray

Abu Shonchoy

Department of Economics

Florida International University

Working Paper 2501

February 2025

11200 SW 8th Street, Miami, Florida 33199

https://economics.fiu.edu/

# Boosting Study Habits with High-Frequency Information: A Field Experiment to Aid Disadvantaged Students*

Tomoki Fujii[a], Christine Ho[a], Rohan Ray[b], and Abu Shonchoy[c]

[a]Singapore Management University
[b]National University of Singapore
[c]Florida International University

**Abstract**

Extended school closures during the COVID-19 pandemic disrupted students' study habits and routine educational engagement, especially in low-income settings where distance education often fails to reach disadvantaged populations. We use a field experiment in rural Bangladesh to determine whether increasing parental engagement can mitigate these disruptions, particularly in the post-pandemic recovery stage. Our findings reveal that a high-frequency information intervention—delivered through weekly text messages and automated voice calls—significantly increases parents' awareness and children's self-study hours, particularly in households lacking access to technology. By disseminating information on available learning resources, teachers' contact details, and the benefits of education, the intervention boosts daily self-study hours by 15 percent. Although Bangladesh's simplified post-pandemic school promotion and shortened syllabus constrained our ability to measure academic improvements, the intervention narrowed study-hour inequalities, promoting upward mobility (and reducing downward mobility) among households without technology access. Shapley-value decomposition analyses indicate that 5-20 percent of the reduced inequality is attributable to the direct treatment effect. Better parental involvement—encouraging children to use learning resources and more household investment in private tutoring—appears to be an important causal channel. Our findings underscore the potential of scalable, low-cost, parent-focused programs to bolster learning continuity under adverse conditions — particularly important for low- and middle-income countries.

**Keywords:** high-frequency information, study hours, post-pandemic recovery
**JEL:** D91, H75, I24, I25, O15

*We are grateful to Akib Khan and Saravana Ravindran for their invaluable comments and feedback. We also thank MOMODa Foundation for data collection support, Metakave for technical assistance, and Rezwanul Haque for research assistance. Our sincere gratitude goes to the participating schools, students, and parents. The trial is registered in the AEA RCT registry (AEARCTR-0007674). This study focuses on students' study hours—one of the primary outcomes in our Pre-Analysis Plan—and incorporates additional non-experimental analyses on how our intervention affects inequality in study hours. The research is approved by SMU-IRB (IRB-21-064-A059(421)) and FIU-IRB (IRB-21-0164). Funding is provided by the Abdul Latif Jameel Poverty Action Lab (J-PAL) Post-Primary Education Initiative (Education Research and Scale for COVID-19 Recovery Request for Proposal) (PPE-1800). All errors remain our own.

# 1 Introduction

A widely cited principle in educational research is the importance of "time on task" in influencing student achievement (Asadullah et al., 2021; Hattie, 2008). The amount of time students actively engage in learning—encompassing quality instructional time and structured home study hours, including self-study and homework—can positively influence their academic trajectories (UNESCO, 2016; World Bank, 2016). However, extended school closures during the COVID-19 pandemic had introduced unprecedented challenges to students. This was particularly evident in resource-constrained countries, where technologies to compensate for the loss of traditional face-to-face education did not reach those who needed them the most. Consequently, many students missed critical school instructional hours, which disrupted routine self-study activities (Amin et al., 2021; Baird et al., 2020).

For example, Bangladesh had faced pandemic-related school closures for 1.5 years since March 2020. During this period, the Bangladesh Ministry of Education adopted a multi-modal strategy to deliver educational programs to students through both online (e.g., YouTube and Facebook) and offline (e.g., television programs) platforms (Biswas et al., 2020; Sarwar et al., 2020).[1] However, the actual use of these resources was highly limited, potentially due to the lack of information, motivation, or supervision (Asadullah et al., 2020). Studies conducted during the pandemic find increased time spent on household chores (Baird et al., 2020) and a considerable reduction in study hours (71 percent decrease in daily study hours for girls compared to pre-pandemic levels) (Amin et al., 2021). These concerning findings are linked with greater learning loss, lower grade progression, and potentially lifetime income consequences (Donnelly and Patrinos, 2021).

How can we help students enhance their self-study habits, particularly for disadvantaged students facing widening educational inequalities? This question is particularly salient in the post-pandemic recovery period, as schools reopened and resumed regular academic activities. One promising approach is to position parents at the forefront of academic oversight. Evidence from a randomized controlled trial (RCT) by Fujii et al. (2023) in rural Bangladesh highlights the effectiveness of high-frequency information (HFI) centered on school attendance, which significantly boosted students' academic engagement before the pandemic in 2017-19. Through weekly SMS and automated voice calls to parents of at-risk secondary school students, the intervention encouraged greater parental

[1] Live Television Programs were hosted by Sangsad TV through national TV broadcasting, where academic content relevant to each grade was telecasted every weekday at pre-announced time slots. Konnect.com is an additional government-hosted online platform offering video and PDF resources. Appendix Figure A1 illustrates some of these educational resources. There were also attempts to host academic and co-curricular classes for students and live sessions for parents and teachers on different platforms, and students were encouraged to stay in constant communication with their teachers to address study-related queries. These various initiatives, mainly through the use of digital platforms as well as offline sources, were designed to mitigate learning loss during extended periods of school closure (Datta, 2022; Khan et al., 2021; UNICEF, 2020).

investment in children. Inspired by these findings, we designed a novel HFI program covering four key information domains: offline TV education programs, online educational resources, teachers' contact details, and the long-term benefits of continuing education.

We conducted this study among 1,200 secondary school students in Gaibandha—a rural district in northern Bangladesh—during three key phases of the COVID-19 pandemic: the school closure, the partial (soft) reopening, and the resumption of regular academic activities. At the baseline, our data show that 60 percent of the sampled students lived in households with television, yet only 39 percent regularly watched educational telecasts. Meanwhile, although all the households had mobile phones, just 33 percent had internet access on those phones, and only a tiny fraction of students had ever viewed educational content online. We find that students spent relatively few hours on self-study, a finding consistent with Amin et al. (2021). During the baseline, we further found that parents were largely unaware of existing educational content across various media platforms. Only 33 percent of students' parents knew about TV-based education programs, and just 9 percent were aware of online platforms offering similar resources. These low levels of awareness are concerning, given that parents serve as a principal source of study-related support for students—98 percent of the students surveyed reported receiving some help from their parents over the past year.

With regular schooling being interrupted and limited student-teacher interactions, parents shoulder a greater responsibility for their children's education. Although parental involvement is high, many parents are either uneducated or lack adequate information, making them less able to offer meaningful guidance or supervision. This issue is particularly pressing in Gaibandha, our study region, where adult literacy stands at only 38 percent. In such contexts, reliance on parents during school interruptions can exacerbate educational inequalities, as students with less educated parents may not receive the constructive help they need. Even if illiterate parents cannot replace teachers, they can still play a critical role by directing children towards the right resources and monitoring their self-study habits, provided they have the necessary information. It is therefore essential to ensure that parents are aware of publicly available educational programs and resources to support, guide, and supervise their children effectively.

To address these challenges, we implemented a HFI intervention aimed at encouraging parents to actively monitor and support their children's study habits. The intervention was evaluated through a randomized controlled trial (RCT), in which half of the student-parent pairs were assigned to receive the HFI program (treatment group), while the other half followed their usual routines without additional support (control group). The intervention primarily sought to close informational gaps by increasing parents' awareness of available educational resources. A secondary objective was to foster greater parental engagement, while motivating families to prioritize their children's continued education during a period of considerable disruptions.

We find better educational activity in the treated household, as captured through self-study hours. Our intent-to-treat (ITT) estimate shows that students in the treated households spent 0.53 hours more on self-study-related activities compared to the control mean of 3.53 hours per day (15.01 percent increase), which is highly statistically significant. This impact is observed consistently across all sub-samples: boys only, girls only, households with and without access to digital technology, and children of more and less educated parents. However, we observe statistically significant impact heterogeneity only by household access to technology (TV or internet). Specifically, compared to households with such access, households without access to technology experience substantially greater increases in study hours ($p$-value = 0.004). Treated students without [with] access to technology studied 0.84 [0.39] hours more per day on average than their counterparts in the control group. These findings suggest that the intervention may be particularly effective in contexts where alternative sources of educational content or information are limited, helping to reduce education inequality faced by disadvantaged students. Furthermore, while the estimated effect on study hours for girls is larger than that for boys, the difference is not statistically significant ($p$-value = 0.110).

We further examine mobility in average daily study hours, where upward mobility and downward mobility, respectively, refer to an increase and decrease in a student's relative position in the study hour distribution compared with the control group. We find that the intervention led to significant improvement with a 13.53 percentage point improvement in upward mobility ($p$-value = 0.000) and a 13.36 percentage point reduction in downward mobility ($p$-value = 0.000), and the results remain similar even when we examine the changes in quantiles beyond different thresholds. These effects were particularly strong for females, households without technology access, and households where no parent had attended school, highlighting the potential of such interventions to address structural inequalities in education access.

Unfortunately, we lacked convincing instruments to measure the learning effects of the enhanced study hours as a downstream impact. First, following COVID-19 disruptions, Bangladesh's education authorities shortened the secondary school syllabus in 2022 to emphasize essential material while also adopting a more flexible "promoted/not promoted" grading system. This policy shift replaced rigorous examinations with simplified assessments, limiting our ability to evaluate performance gains. Second, we administered a brief, short mathematics test with ten questions—originally designed for baseline balance checks—at both baseline and endline. The results showed no significant effect on test scores, either for the overall sample or within any subgroup. This null finding could reflect the test's brevity, its unincentivized format, or the generally slow pace at which learning outcomes tend to manifest (Brisson et al., 2017). Nevertheless, improving study habits in the post-pandemic recovery period is an important first step towards improving learning in the long run.

Our main mechanism, as hypothesized, is greater parental awareness. We find that HFI intervention significantly increased parental awareness of educational resources among treated households. This impact remains robust in different sub-samples, as mentioned before. However, we find that households lacking access to television or internet exhibit substantially greater increases in resource awareness compared to those with such access ($p$-value = 0.013). Additionally, parents in the treatment group were significantly more likely to guide their children towards both online and offline educational programs for study support, an effect observed consistently across the entire sample and all sub-samples, with no evidence of heterogeneity. We also document a statistically significant increase in resource allocation for private tutoring, particularly benefiting treated girls and households with at least one parent who had attended school. Nevertheless, we find no significant heterogeneity across other observable attributes.

To assess how the treatment effect contributed to the observed reduction in study-hour inequality, we employ a regression-based decomposition using the Shapley-value approach. To our knowledge, this is the first study to conduct a regression-based inequality decomposition analysis in an experimental setting. This method enables us to decompose changes in the Gini coefficient of study hours—a measure of inequality that captures how study time is distributed among students—into components attributable to the treatment effect and broader time trends. Our results indicate that, depending on the chosen inequality metric, 5–20 percent of the decline in study-hour inequality can be directly attributed to the intervention. Similar patterns were observed when applying alternative inequality measures, reinforcing the robustness of these findings.

The intervention implemented in this study is automated, leverages readily available low-cost mobile technology, and draws on publicly accessible educational resources. In our study region, 91 percent of households own mobile phones, and sending an SMS or making a one-minute voice call costs roughly \$0.01, enabling high compliance rates (i.e., high receipt of text messages and voice calls). This approach offers a cost-effective means of closing information gaps and empowering parents to play a pivotal role in their children's education. Because it efficiently employs existing technology and educational resources, the intervention is rapidly scalable and can be adopted in other low- and middle-income countries in Asia and Africa with similar infrastructure.[2]

Our intervention is the first study aimed at boosting study habits of students in the post pandemic recovery phase and help them return to their pre-pandemic educational trajectories. Extended periods of school closure during the pandemic impeded consistent study routines, leading to significant learning losses and difficulty in re-establishing

[2] Although mobile phone penetration is high, most households own only basic feature phones, particularly in rural areas; SMS and voice calls still work effectively on these devices. The number of smartphone users in Bangladesh is expected to grow, but only 41 percent of mobile phone users had smartphones as of 2020 (Bhuiyan, Mohsin, 2021). Similar trends can be found in other developing and less developed countries worldwide (Cheney, Catherine, 2018; Silver and Johnson, 2018).

routine educational activities when schools reopened (Betthäuser et al., 2023). Existing strand of literature has focused on learning losses and the resulting economic impacts (Engzell et al., 2021; Hanushek and Woessmann, 2020; Makino et al., 2021), while other studies aimed to find ways to alleviate learning loss through remedial online or mobile lessons and telementoring (Angrist et al., 2020; Clark et al., 2021; Hassan et al., 2023). Our research departs from these studies and looks at reducing pervasive information gaps about publicly available remedial educational contents that are already in place, and thereby improve study habits in the post pandemic period when schools were beginning to operate but not at full capacity. Thus, we offer an affordable way of exploiting existing resources to help children from disadvantaged backgrounds to reestablish previous study routines and potentially recuperate from learning losses.

Second, our study informs targeted policy design and advances understanding of how to reduce educational disparities, especially in a post-pandemic recovery context. The pandemic not only disrupted study habits but also amplified pre-existing disparities in educational outcomes between students from marginalized households and those from better socioeconomic backgrounds. Children from wealthier families often had greater access to digital technologies, private tutoring, and conducive learning environments, enabling them to better adapt to remote learning (Di Pietro et al., 2020). In contrast, students from disadvantaged households faced significant barriers, including lack of internet access, fewer parental resources for academic support, and increased economic pressures that diverted attention away from education (OECD, 2021). Addressing such inequalities has become a critical focus in educational policy and research. Previous interventions targeting marginalized students have primarily focused on expanding access to educational resources and reinforcing parental engagement to help overcome these challenges (Alon, 2007; Castleman and Page, 2015; Fack and Grenet, 2015; Lai et al., 2015). However, the heterogeneity of outcomes based on gender, access to technology, and parental education remains underexplored. Our study directly addresses this gap by analyzing the differential effects of information nudges on these dimensions. For instance, we assess whether access to a television or the internet mediates the impact of our interventions or whether parental education levels influence the efficacy of these nudges.

Finally, our paper contributes to the literature on closing parental information gaps regarding their children's education through various channels, including parent-teacher meetings, emails, phone calls, and text messages (Avvisati et al., 2014; Barrera-Osorio et al., 2020a,b; Bergman, 2021; Bergman and Chan, 2021; Berlinski et al., 2021; De Walque and Valente, 2018; Dizon-Ross, 2019; Kraft and Dougherty, 2013; Kraft and Rogers, 2015; Rogers and Feller, 2018; York et al., 2019). Previous studies in this area largely consist of interventions such as sending parents individualized updates on students' behavior, attendance, and test scores, or providing information on the benefits of education, and have mainly been carried out in developed countries. Our study differs from the existing literature in terms of culture, sophistication of widely available technologies, and average income and educational levels of the household, among others. Thus, our paper adds to a relatively thin evidence base on the importance of parents' role in their children's education in an economically marginalized and challenging rural setting in South Asia.

The rest of the paper is organized as follows. Section 2 describes the context and the study design, including a discourse on data, power calculations, randomization balance checks, and HFI compliance rates. Section 3 presents the main identification strategy, and Section 4 presents the results. Section 5 performs a decomposition analysis and Section 6 concludes.

# 2 Research Design

## 2.1 Background

Bangladesh ranks third among the list of countries that experienced the highest number of days of full school closure since the onset of the pandemic (UNICEF et al., 2021). Schools were completely shut down between March 2020 and September 2021. Though schools started to operate again in September 2021, they were yet to function at full capacity (Mahmud, 2021). Schools were open to students one day a week per grade for the better half of 2022, suggesting that students were still experiencing considerable disruptions in their studies on top of the school closure for one and a half years. Such pandemic-driven interruptions compromised the future of an entire generation of school-age children and put them at a greater risk of education discontinuation. It was, therefore, imperative to encourage students to get back to their routine educational engagements in the post-pandemic phase and motivate them to continue with their studies.

Our intervention took place in a relatively poor agricultural district, Gaibandha, located in northern Bangladesh. Gaibandha is a rural district where educational attainment has been historically low. It ranked $57^{th}$ in the literacy rate among the 64 districts in the country, with 38 percent of the adult population being literate, which is lower than the national average of 51 percent (World Bank, 2020). Similarly, 48 percent of the population were below the poverty line in Gaibandha, which is much higher than the national average of 32 percent (World Bank, 2020). Agriculture is the primary occupation for around 71 percent of the working population in this district. Since the opportunity cost of attending school might be relatively high for children from agricultural households, the district of Gaibandha presents a setting where children may be highly susceptible to dropping out of school due to full school closure.

## 2.2 Experimental Design

We first collected the detailed student roaster from the four participating secondary educational institutes in our study area after getting the consent from the school headmaster. We randomly kept only one child per household in the roster in order to maintain analytical tractability and to avoid the possibility of within-household spillovers and confusion about the messages sent to the households.

From this roaster, we selected a total of 1,200 students in grades 6-10 stratified by school, grade, and gender, excluding students from a household without a valid mobile phone. Enumerators then visited households to get parental consent for participation in the study following the pandemic-relevant survey protocol. Upon receiving consent, we completed the baseline data collection and then randomly assigned students to one of the following two experimental arms: households who received information through SMS and automated voice call on a weekly basis (henceforth called the information treatment arm or HFI) and households who received neither SMS nor automated voice calls (henceforth called the control arm).

The information sent to the households in the HFI treatment arm varied every week, and covered the following four types of information:

(I1) When to view grade-appropriate educational programs on television.

(I2) How to access educational programs on different online platforms such as YouTube.

(I3) Contact details of the class and subject teachers for English and Mathematics.

(I4) General information on the benefits of continuing education.

Our intervention was aimed at increasing the awareness of televised and online educational resources for parents, improving teacher involvement in children's education, and helping parents realize the benefits of more years of schooling. Some of the messages were customized to suit the needs of each study participant. That is, the exact message varied by the grade of the student for (I1) and (I2) and by the grade-section combination for (I3). For (I4), all treatment students received the same message. As discussed in Appendix A.1, the minimum detectable effect size of this study is 0.12 standard deviation with statistical significance and power of 0.1 and 0.8, respectively, suggesting that our study is likely to capture moderately large effects.

We carefully structured our RCT to ensure that no information was withheld from the control group, and instead we aim to reinforce only publicly available information for the HFI treatment arm. Table A1 in Appendix A provides examples of the SMS and voice calls sent to households in the information treatment arm.

The baseline survey took place between September and October 2021, after which participants were randomly assigned to the information treatment or a control group.

The HFI intervention started in November 2021 and continued until December 2022, spanning the period of school closures due to the Omicron surge, as well as the gradual reopening and eventual return to regular schooling.

A short midline survey was conducted in June–August 2022, followed by an endline survey in January–February 2023. We also conducted a short follow-up survey in September 2023, about eight months after the HFI intervention ended. Figure 1 illustrates the timeline of our study.

These surveys collected data on study hours—our primary outcome—and additional household-level information, including the availability of digital technology and parental education. Unfortunately, we could not objectively measure the hours or minutes of educational TV and online viewership, as doing so would require technological sophistication (e.g., account-based subscriptions) that were not accessible to us. Even if such data were available, it would be difficult to ascertain who within the household actually watched the content. Moreover, participants could view education programs through neighbors' televisions or access online resources using friends' and neighbors' devices, further complicating any attempt at precise measurement.

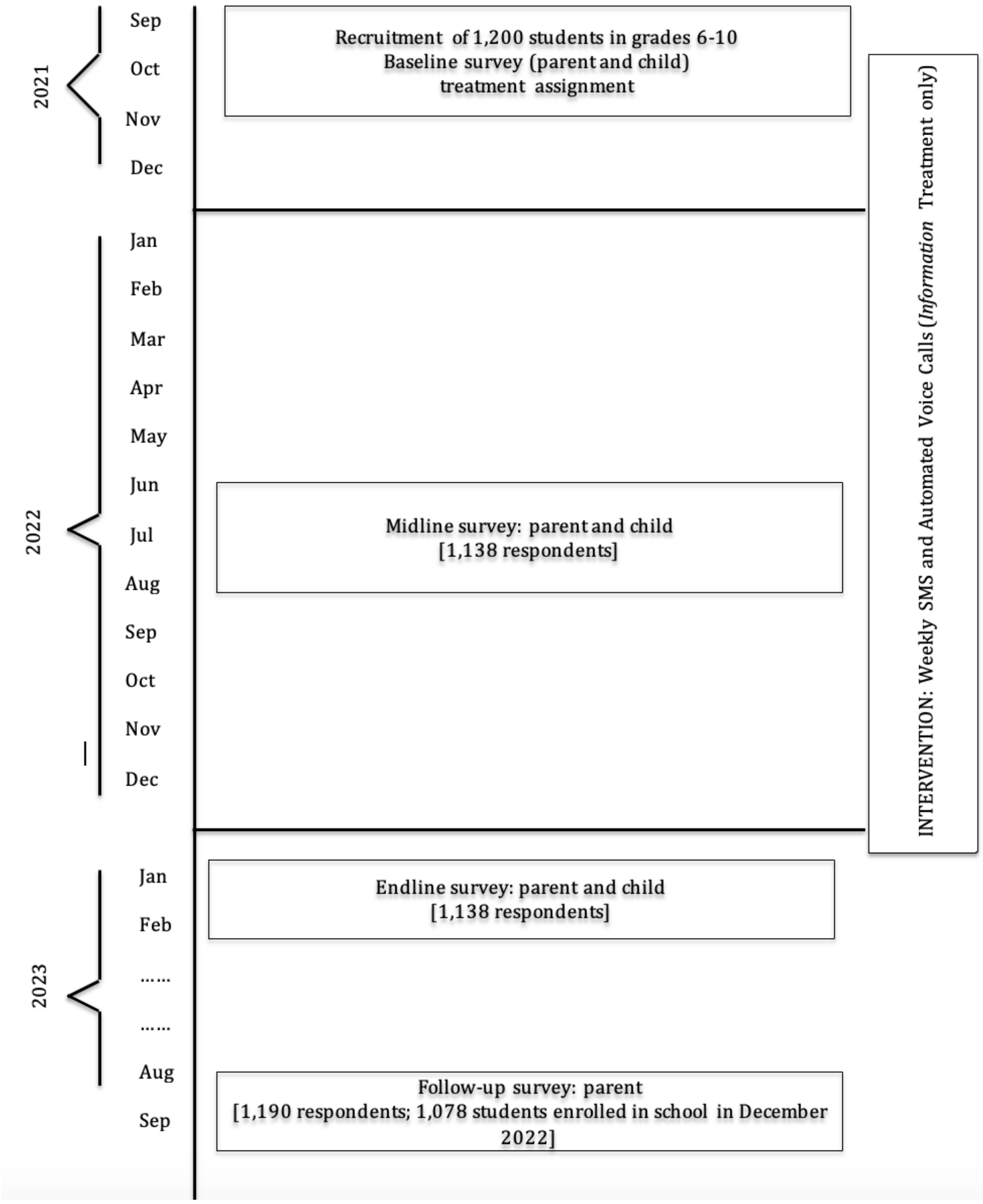
## 2.3 Treatment Compliance

It is important that the compliance rates for SMS and voice calls (i.e., receipt of SMS and voice calls) are sufficiently high among those treated for there to be any impact on study behavior. As can be seen from Table A2 of Appendix A, almost all households in the information treatment arm received SMS in most of the first 30 weeks of the intervention. However, we acknowledge that the successful delivery of an SMS does not necessarily imply that the households read it.

Another relevant measure of treatment compliance is the number of voice calls picked up by a household. On average, 76 percent of treated households in the information treatment arm received and listened to a voice call. Thus, the average HFI compliance rate was relatively high for our intervention. A breakdown for the reception of SMS and voice calls by week is given in Table A2.[3]

[3] The imperfect compliance could have occurred for a variety of reasons such as invalid phone numbers (because the phone number had changed since the time of recruitment, for example) or the phone being switched off (for an extended time for the case of SMS). As Table A2 shows, there were notable drops in compliance rates in Weeks 10, 20, 23, and 24 for SMS and Weeks 3, 4, and 12 for voice calls. These were due to implementation issues such as system-wide failure in mobile network connectivity.

Figure 1: Timeline of the Study




Note: The academic year coincides with the calendar year for grades 6-9.

## 2.4 Measurement of the key outcomes

### Measurement of Study Hours

Our primary outcome of interest is study hours. A key challenge in this analysis is determining the most reliable measure of study hours, as the numbers reported by children and parents can differ in availability, accuracy, and relevance. Students self-reported their study hours at the baseline and endline, and parents also reported the students' study hours. Parents additionally reported study hours at follow-up. Understanding which measures better reflect actual study behavior is critical for evaluating the intervention's impact on educational engagement.

As shown in Table A3 in Appendix A, the correlations between student-reported and parent-reported study hours are high at 0.521 and 0.439 for the baseline and endline, respectively, suggesting that they report coherent study hours. However, the levels of study hours are different, as shown in Table 1. We prefer to use the parent-reported measure for the following three reasons.

First, the parent-reported study hours provide a more plausible account of children's study behavior than the student-reported measure. At the baseline, parents and students reported study hours for August, which typically experiences reduced academic activity, given the onset of the post half-year exam period (Table 1). The study hours at the endline is for November and characterized by heightened academic activity due to forthcoming high-stake grade progressing final examinations in December. Interestingly, the student-reported study hours show no significant difference between baseline and endline, which is at odds with expected seasonal variations in study behavior. In contrast, parent-reported study hours are nearly double at endline compared to baseline, aligning with the increased academic demand during the examination period. This pattern highlights the plausibility of the parent-reported data to be more sensitive to contextual changes in educational engagement.

Second, the validity of the parent-reported study hours is further supported by their stronger correlation with key educational outcomes. At the baseline, parent-reported study hours exhibit a higher correlation with students' test scores and motivation to study compared to student-reported study hours (Table A3 in Appendix A). This indicates that parent-reported data better capture the components of study behavior that translate into improved learning outcomes and intrinsic motivation. These findings reinforce the decision to use parent-reported study hours as the primary measure in this analysis.[4]

Finally, students are arguably more likely to misreport the study hours, as they may believe that their reported study hours could somehow be conveyed to their parents or

[4]On the other hand, at the endline, the parent-reported study hours are not correlated well with other measures, except for the student-reported study hours. The parents may be generally busier during the harvesting time in November, so the measurement errors tend to be larger. Partly for this reason, we mostly examine the outcome of the follow-up survey.

classmates. On the one hand, students may over-report their study hours so that their parents may feel better. On the other, if they believe that classmates adjust towards the mean of the reported study hours, students may under-report their study hours to gain competitive advantage. The former effect can be large at the baseline, whereas the latter effect may be prominent at the endline. Hence, the parent-reported measures will likely offer a more reliable assessment of students' educational engagement.

**Measuring Mobility**

While the level of study hours is of interest, the gap in the study behavior among students is also of interest. Even if our intervention increases the study hours, the implications of our intervention for inequality will depend on whose study hours have increased by our intervention. Since our intervention intends to help disadvantaged students study harder by addressing the information gap, we expect to see higher upward mobility and lower downward mobility relative to the distribution of the study hours for the control group. Thus, to measure the mobility of students in the distribution of study hours, we first find the quantiles of study hours for each individual in the control-group distribution at the time of observation.

To this end, we sort the control observations from the lowest to the highest regarding the average number of study hours per day.[5] Without loss of generality, we can let $y_1^C \leq y_2^C \leq \cdots y_{N_C}^C$, where $y_i^C$ is the $i$th smallest observation of daily study hours in the control group and $N_C$ is the number of control observations. If $y_i$ has no ties, we assign the value of quantile $q_i = (i - 1/2)/N$. When $y_i$ has $j(> 1)$ ties (i.e., there are $j$ observations of $y_i$ in the control group), we take the average of $(i - 1/2)/N_C, (i + 1 - 1/2)/N_C, \cdots, (i + j - 1 - 1/2)/N_C$ and assign $q_i = (i + j/2 - 1)/N_C$ for the quantile of $y_i$.

Next, we map each treatment observation to the quantile in the control distribution. If a given treatment observation $y^T$ has a corresponding observation $i$ in the control group such that $y^T = y_i^C$, we map this treatment observation to $q_i$. If a treatment observation is smaller than $y_1^C$ [greater than $y_{N_c}^C$], we assign a quantile value of $q = 0$ [$q = 1$]. Finally, if a given treatment observation $y^C$ satisfies $y_i^C < y^T < y_{i+1}^C$ for some $i$ in the control observation, we linearly interpolate and assign $(i - 1/2)/N + (y^T - y_i^C)/(y_{i+1}^C - y_i^C)$ as the quantile value for this observation. Based on this process, we assign a quantile in the control distribution to each $y_i^g$ for treatment group $g = \{C, T\}$ and individual $i$.

Therefore, for each individual in control and treatment groups, we have the quantile of study hours for each of the baseline, endline, and follow-up surveys, which we denote by $q_{it}$, where $t = 0$ and $t = 1$ respectively represent the baseline and comparison periods, where the latter of which may be the endline or follow-up. We define upward mobility

[5]We ask questions about the daily study hours for typical weekdays and weekends. We give a weight of 5/7 and 2/7 to the former and latter, respectively, to obtain the daily average study hours.

Table 1: Summary of Self-Reported Study Hours by Parents and Students

| Respondent | Recall Span | Reference period | Mean | (s.d.) |
|---|---|---|---|---|
| *Baseline (conducted in Sep 2021)* | | | | |
| Parent | One month | Aug 2021 | 2.29 | (1.27) |
| Student | One month | Aug 2021 | 2.75 | (1.29) |
| *Endline (conducted in Jan 2023)* | | | | |
| Parent | One month | Nov 2022 | 5.37 | (2.78) |
| Student | One month | Nov 2022 | 2.97 | (1.58) |
| *Follow-up (conducted in Sep 2023)* | | | | |
| Parent | One year | Sep 2022–Aug 2023 | 3.79 | (1.35) |

Note: The above table shows the average study hours, as reported by parents and students at the baseline, endline, and follow-up surveys. The baseline survey began in September 2021, and respondents were asked about study hours in August. The endline survey was conducted in January, 2023, and respondents were asked about study hours in November, 2022. The follow-up survey was conducted in September of 2023, and parents were asked about study hours in the past year.

and downward mobility as follows

$$UM_i^0 = \mathbf{1}_{(q_{i1} > q_{i0})}, \quad DM_i^0 = \mathbf{1}_{(q_{i1} < q_{i0})}, \tag{1}$$

Because we have ties and measurement errors, we may be interested in large enough changes. Therefore, we also consider the following upward and downward mobility measures, where $a > 0$ is a given constant:

$$UM_i^a = \mathbf{1}_{(q_{i1} > q_{i0} + a)}, \quad DM_i^a = \mathbf{1}_{(q_{i1} < q_{i0} - a)}, \tag{2}$$

We use $UM^a$ and $DM^a$ for $a \geq 0$ as mobility measures for our purpose.

## 2.5 Randomization Balance

The control and treatment groups should have the same expectation for both observable and unobservable characteristics due to the randomization. To check for the potential violation of randomization, we conduct a $t$-test of equality of means for several individual and household-level observable baseline characteristics such as study hours, gender of the student, education level of household head, availability of television or internet at home, number of household members, number of male members, number of female members, and ownership of household assets such as agricultural land, radio, television, bank account, NGO account, and bicycle, and baseline test score. As Table 2 shows, nearly all the observable characteristics and, more importantly, study hours are similar between the control and treatment groups. A notable exception is bicycle ownership, which is controlled for in our empirical analysis as detailed in Section 3. Therefore, there is no evidence that our randomization was compromised.

Table 2: Summary Statistics and Balance Check

| | Control (1) | Treatment (2) | Overall (3) | Orthogonality† (4) |
|---|---|---|---|---|
| Participating child is female | 0.568 | 0.568 | 0.568 | 1.000 |
| Daily study hours | 2.254 | 2.324 | 2.289 | 0.340 |
| Number of household members | 4.547 | 4.513 | 4.530 | 0.643 |
| Number of male household members | 2.188 | 2.137 | 2.162 | 0.365 |
| Number of female household members | 2.358 | 2.377 | 2.368 | 0.757 |
| Household head has no schooling | 0.350 | 0.375 | 0.363 | 0.368 |
| Household head has primary education | 0.297 | 0.298 | 0.298 | 0.950 |
| Household head has secondary education | 0.353 | 0.327 | 0.340 | 0.330 |
| Household has agricultural land | 0.990 | 0.993 | 0.992 | 0.526 |
| Household has radio | 0.020 | 0.017 | 0.018 | 0.667 |
| Household has television | 0.605 | 0.603 | 0.604 | 0.953 |
| Household has bank account | 0.235 | 0.227 | 0.231 | 0.732 |
| Household has NGO account | 0.467 | 0.438 | 0.452 | 0.325 |
| Household has television/internet at home | 0.703 | 0.690 | 0.697 | 0.616 |
| Household has bicycle | 0.537 | 0.475 | 0.506 | 0.033 |
| Math score at baseline | 5.228 | 5.205 | 5.217 | 0.853 |
| Observations | 600 | 600 | 1,200 | 1,200 |

Note: Column (1) shows the mean value of the variables for the control arm and Column (2) shows the mean value of the variables for the information treatment arm. Ownership of assets (agricultural land, radio, television, bank account, NGO account, access to television/internet, bicycle) is a binary variable that takes a value of 1 if the household owns the asset, and 0 otherwise. "Math score at baseline" shows the raw score in a mathematics test, comprising of ten questions, administered by the research team at baseline. Column (3) shows the mean value for each variable. Column (4) shows the $p$-value for joint orthogonality.

# 3 Empirical Strategy

We use a simple intention-to-treat analysis to test the impact of our intervention on our main outcome of interest—study hours. We denote the outcome of interest for student $i$ at time $t$ by $y_{it}$, where $t = 0$ and $t = 1$ respectively denote the baseline and follow-up surveys. As we are most interested in the persistent effect of the treatment, we primarily use the follow-up survey for $t = 1$, but we also consider the endline survey for $t = 1$ as an alternative.

We denote the unbalanced covariate at the baseline (ownership of bicycle) by $\mathrm{X}_{i0}$ and the strata-specific fixed effect by $\mu_{s(i)}$, where $s(\cdot)$ is a mapping from individual to the stratum that the individual belongs to. $T_i$ denotes the treatment assignment indicator, which takes unity if individual $i$ is assigned to the treatment group and zero otherwise. Using these notations, our primary specification is as follows:

$$y_{i1} = \alpha_0 + \alpha_1 T_i + \alpha_2 \mathrm{X}_{i0} + \mu_{s(i)} + \epsilon_{i1}, \tag{3}$$

where $\alpha_1$ is the coefficient of primary interest. We will report standard errors clustered at the strata level. We use the same econometric specification to study the impact of our intervention on upward and downward mobility, and other outcomes. Additionally, we use an ANCOVA specification to control for the baseline measure of the outcome variables of interest. The results are qualitatively and quantitatively similar and are available upon request.[6] Finally, we present the Benjamini-Krieger-Yekutieli (BKY) sharpened q-values to account for false discovery rate.

Because our interest is not only in the level of the study hours, but also in whether our intervention can narrow the gap between disadvantaged students and the rest, we examine the impact heterogeneity of our intervention. Specifically, we study the impact heterogeneity along the following three important dimensions—gender (boys vs girls), accessibility to digital technology (households with access to television or internet versus households with access to neither), and education level of parents (parents who attended school versus parents who were never enrolled in school).

As the treatment compliance was not perfect, we also estimate a variant of eq. (3) by instrumental variables regression with $T_i$ replaced with the endogenous compliance indicator. This indicator takes one if the respondent in the follow-up reported that the respondent's household member had received an SMS and zero otherwise, and it is instrumented by $T_i$. $\alpha_1$ in this case represents the treatment effect on the treated.

We also acknowledge that there may be potential spillovers from the treatment to the control group, given that the information we send by SMS can be easily shared with the control group. While this is a possibility, we argue that the scope of such spillovers is likely to be limited during our intervention as the level of interactions among parents is generally limited. If anything, our estimates would represent the lower bound of the actual treatment effect.

**Differential Attrition Between Treatment Arms** The impact estimates could be biased if the attrition rates systematically differ between the information treatment and control arms. While 62 household students discontinued the study by the time of the endline survey, a total of 122 student households attrited at the time of follow-up. Discontinuation could have occurred due to school dropout or migration. We test for systematic attrition between the two experimental arms using eq. (3), and find that the attrition rates are statistically indistinguishable between the two arms (Table A4).

[6] We study the impact of our intervention on reported study hours at the endline as well; however, we do not see any statistically significant effect. This pattern is plausibly explained by the potential measurement error (see footnote 4), time required for households to fully comprehend the intervention and for outcomes such as study hours to change significantly.

# 4 Results

## 4.1 Study Hours

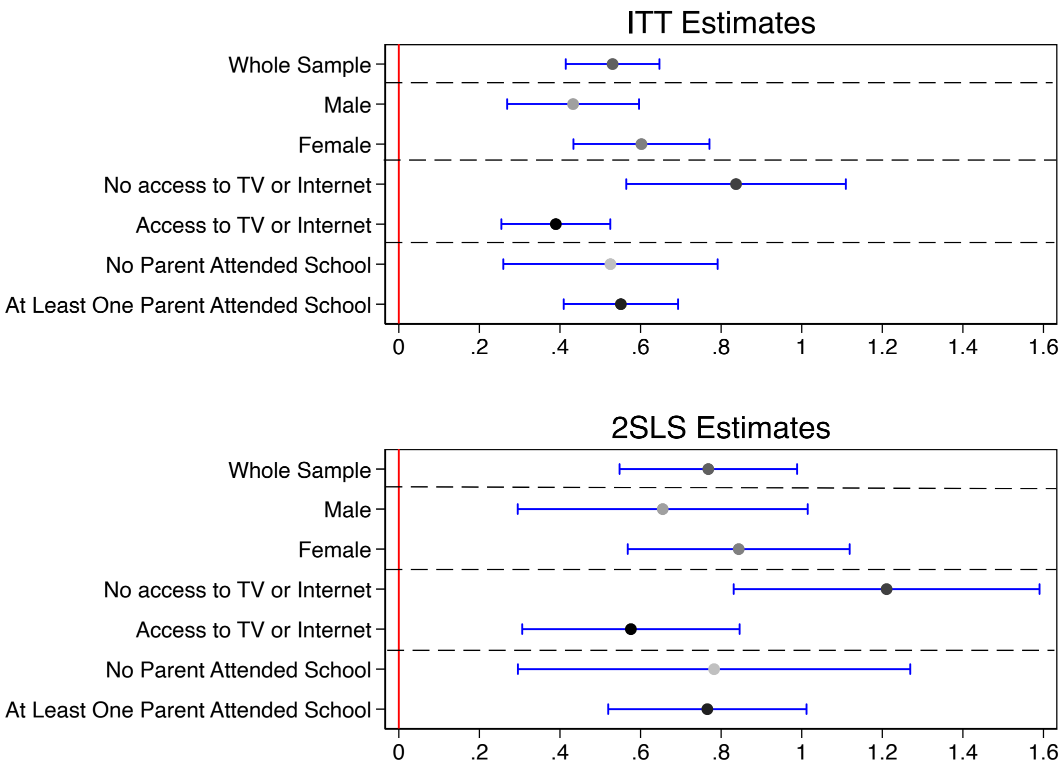
### Level of Study Hours

The intervention resulted in a significant increase in study hours among the treated group, with the effects being observed consistently across all sub-samples (Figure 2; Table A5). While the mean study hours for students in the control arm were 3.53 hours per day, students in the treatment group on average studied 0.53 hours more in the follow-up survey, and this difference is statistically significant at the 1 percent level. However, we observe statistically significant impact heterogeneity only by household access to technology. Specifically, households without access to television or the internet experience substantially greater increases in study hours compared to households with such access ($p$-value = 0.004)—treated students without access to technology studied 0.84 hours more per day on average, whereas the ones without access to such technology studied 0.39 hours more per day compared to their counterparts in the control group. These findings suggest that our intervention is particularly effective when students have limited access to educational information or content.

We further look at the impact of our intervention on the treated. To estimate the treatment-on-the-treated (TOT) effects, we asked the parents in the follow-up survey whether any member in the household received SMS or voice call on education content from our partner organization in the past one year. Though self-reported, all households in the treatment arm reported receiving the intervention, while none in the control arm reported the same, indicating perfect compliance. Leveraging this compliance, we employ a two-stage least squares (2SLS) regression, instrumenting compliance with treatment assignment. The 2SLS estimates reveal significant increases in study hours across the entire sample as well as in various subgroups. These TOT estimates are larger than the intention-to-treat (ITT) effects, as expected, since they reflect the impact of the intervention on the subset of households that complied with the treatment.

It is also worth noting that our intervention appears to be gender-neutral. While the estimated effect on study hours is somewhat larger for girls than for boys, the difference is not statistically significant ($p$-value = 0.110). We also find no significant difference in treatment impacts by parental schooling. These findings suggest that closing gaps in study behavior stems more from differences in technology access than from factors like gender or parental educational background. Consequently, this has important implications for how study hours may evolve—an issue we now turn to.

Figure 2: Impact of Treatment on Study Hours




Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable in the above figure is a continuous variable measuring daily study hours in the past one year. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

### Mobility in Study Hours

We start with the $UM^0$ (upward mobility) and $DM^0$ (downward mobility) defined above. Our intervention led to a significant upward mobility in study hours for the treated sample, alongside a marked reduction in downward mobility (Figure 3; Table A6). More specifically, we observe a significant 13.53 percentage point improvement in upward mobility ($p$-value = 0.000) and a 13.36 percentage point reduction in downward mobility ($p$-value = 0.000) for students in treated households. These effects were particularly strong for females, households without access to television or internet, and households where no parent had attended school, highlighting the potential of such interventions to address structural disparities in education access. Because Figure 3 does not consider the changes in quantiles, we also vary the value of $a$ and rerun similar regressions, dropping the observations that are within the share $a$ among the top [bottom] of the control distribution at the baseline. We drop these observations because $UM^a$ [$DM^a$] is equal to zero for such observations by definition. As shown in Figure A2 in the Appendix, we see that the treatment significantly increases upper mobility and decreases downward mobility for a wide range of values of $a$.

Heterogeneity analyses reveal that the impacts on both upward and downward mobility are larger in absolute value for traditionally more disadvantaged groups of students, such as females (relative to males), those without access to technology (relative to those with access), and children with parents who did not attend school (relative to children with parents who attended school), even though the difference is statistically insignificant. Hence, our information nudges can compensate for the challenges that disadvantaged students face and offer a low-cost and scalable solution to bridge the gaps in study behavior and possibly learning in resource-poor settings.

## 4.2 Learning Outcomes

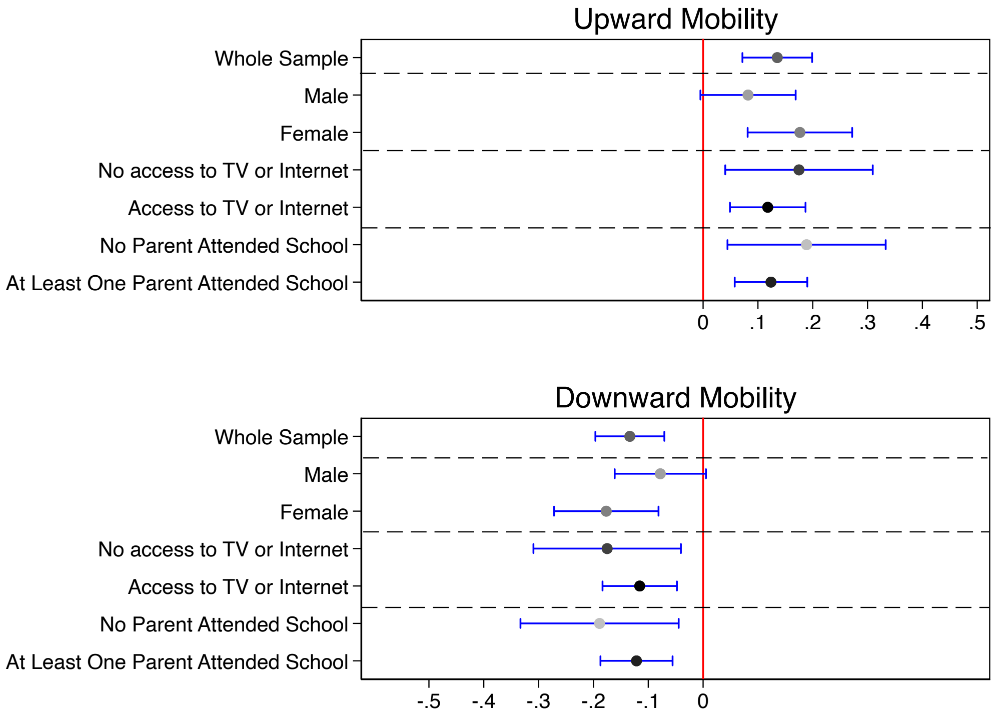
Our intervention, which significantly increased study hours, holds promise for improving learning outcomes. Lacking formal academic assessment due to Bangladesh's simplified post-pandemic school promotion and shortened syllabus, we conducted a short mathematics test comprising ten questions at the baseline and endline. Using eq. (3), we studied the impact of information nudges on learning outcomes, as measured by performance in this mathematics test. While we did not observe any impact of the intervention on endline test scores on the overall sample or any sub-sample (Figure A3; Table A7), the potential of the intervention to increase study hours is a promising sign of its effectiveness.

The null effects of information nudges on test scores are unsurprising but underscore the challenges in translating increased awareness into measurable improvements in learning outcomes. While the intervention may have successfully informed households about digital education programs, this increased awareness may be insufficient to overcome more profound structural barriers, such as supply-side issues or a lack of parental and student capacity to engage with the resources effectively. Moreover, the reliance on a short mathematics test to measure learning outcomes may not fully capture broader cognitive and behavioral changes that such interventions could influence. These findings suggest that we cannot expect immediate impacts on learning outcomes from study behavior improvements (Brisson et al., 2017).

## 4.3 Plausible Mechanisms

Let us now explore some potential mechanisms through which the intervention may have impacted our primary outcome of interest—study hours. In particular, we examine parental awareness of different education programs (Figure A4; Table A8) and various sources to which parents directed their children for assistance with study-related matters in the follow-up survey (Figure 4; Table A9).

Figure 3: Impact of Treatment on Mobility




Note: The above estimates are derived from the baseline and follow-up survey conducted with 1,072 parents. The outcome variables in the above figures are mobility measures based on study hours at the baseline and follow-up. "Upward mobility" ["Downward mobility"] are indicator variables that take a value of 1 if the student moved up [down] in the control distribution of average daily study hours between the baseline and follow-up surveys, as defined in eq. (1). The specifications control for unbalanced covariates at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

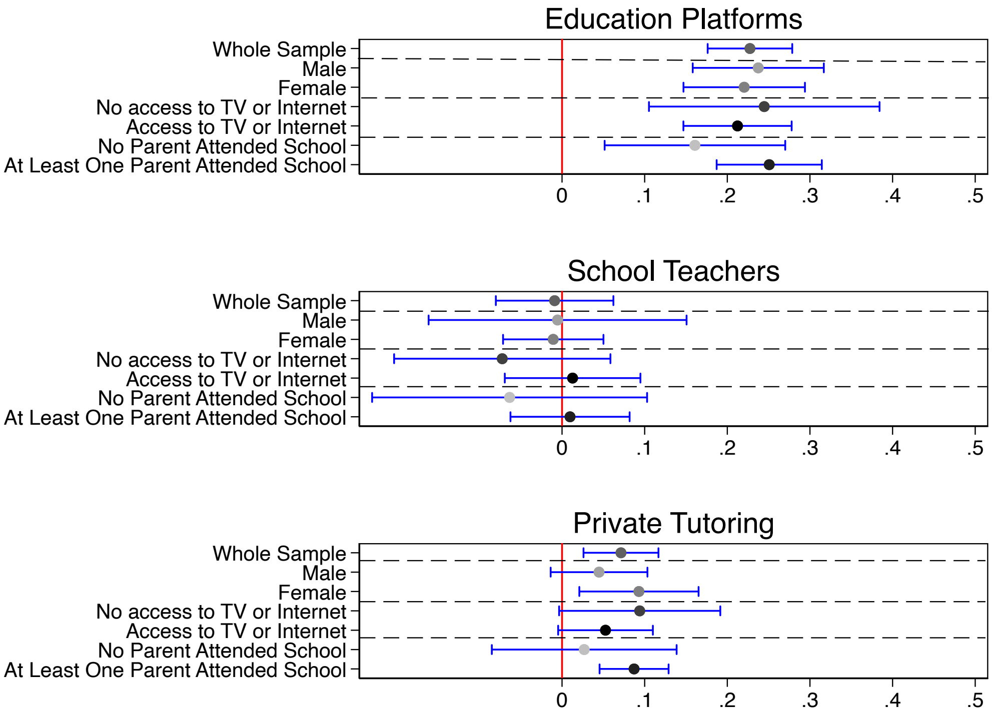
Given that the SMS and voice call interventions were primarily directed to mobile phones owned by parents, increasing parental awareness was a critical prerequisite for influencing changes in children's study hours (Figure A4). The findings indicate a significant increase in awareness of educational programs among treated households, encompassing all sub-samples: boys, girls, households with and without access to digital technology, and children of both more and less educated parents. Further impact heterogeneity across sub-samples reveals that the difference in parental awareness is mostly statistically insignificant. However, one notable exception is access to technology. That is, households lacking access to television and the internet exhibit substantially greater increases in awareness compared to those with such access ($p$-value = 0.013). This heterogeneity is not surprising as those who lack access to technology are less likely to be exposed to the programs that are available on television or the internet. Together with the high compliance rate (Section 2.3), these results show that our intervention has created the intended awareness impacts.

To underscore the importance of awareness, we perform a simple mediation analysis, the awareness of education programs at the follow-up is included as an additional regressor in eq. (3). Figure A5 (Table A10) in the Appendix shows that the awareness of education has a coefficient that is highly significant both statistically and economically regardless of the choice of the sample, and the coefficient on the treatment indicator is generally smaller than those reported in Figure 2. To understand the order of magnitude of the importance of awareness, let us take the estimates for the whole sample. The coefficients for the whole sample in Figures A4 and A5(a) are 0.233 and 0.832, respectively, which suggest that treatment increased the study hours by $0.194 (= 0.233 \times 0.832)$ hours due to increased awareness. Correspondingly, the coefficient on the treatment indicator for the whole sample drops from 0.531 in Figure 2 to 0.336 in Figure A5(b). Thus, about 37 percent (=0.194/0.531) of the total treatment effect on study hours could be attributed to the treatment effect through the awareness channel, and the rest could be attributed to the direct treatment effect. In Section 5, we use a similar specification but extended to two periods with heterogeneity with respect to technology access to quantify the effect of treatment on inequality using a regression-based decomposition approach.

Treated parents were also significantly more likely to guide their children to both online and offline education programs for study-related assistance, an effect observed consistently across the overall sample and all sub-samples, with no significant heterogeneity. This suggests that the intervention effectively shifted parental behavior towards leveraging educational resources, irrespective of individual or household characteristics. However, there was no significant impact on parental guidance-seeking initiatives towards school teachers across any group, indicating that traditional channels of academic support may remain underutilized, perhaps due to limited access during school closures or entrenched barriers in teacher-student engagement.

For household investment in private tutoring, the intervention had a significant positive effect on the overall sample. While we do not observe any significant heterogeneity across genders, technology access, or parental schooling, we found that the information nudge tends to have a larger impact on disadvantaged children as their parents are more likely to direct them to obtain help from private tutors for their study, suggesting that some of the disadvantaged children in the treatment group are benefiting from private tutoring support, who wouldn't otherwise benefit in the absence of the information nudge. Overall, our results underscore the importance of tailoring educational policies to enhance access to formal and informal learning opportunities while addressing structural barriers limiting the effectiveness of school-based support systems.

Figure 4: Mechanism: Sources of Help Directed by Parents




Note: The above estimates are derived from the follow-up survey of 1,077 parents. The outcome variables in the above panel take unity if the parents guide their children to the mentioned source for study-related help in the past year, and zero otherwise. The specifications control for unbalanced covariates at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

# 5 Decomposition analysis

The preceding discussion indicates that our intervention had a positive impact on study hours both in the absolute and relative sense. We have also observed some changes in the behavior. While these results are encouraging, it is unclear how much the treatment effect contributes to the reduction in the inequality in study hours. To address this question, we propose a regression-based decomposition of changes in inequality based on the Shapely-value approach.

The fundamental idea behind the proposed approach is somewhat similar to the Oaxaca-Blinder decomposition after Oaxaca (1973) and Blinder (1973). It is a regression-based decomposition approach that has been widely used to describe the wage gaps across different demographic groups, among others. As with the Oaxaca-Blinder decomposition, we consider the marginal contribution of a particular variable by letting it change while fixing all other variables constant. However, unlike the Oaxaca-Blinder literature, we decompose an overall inequality measure rather than a specific gap.

There are also regression-based studies that attempt to decompose inequality by essentially computing a series of marginal contributions, but they tend to impose an arbitrary sequence of variables that are allowed to change (Juhn et al., 1993; Yun, 2006). To address this issue, we adopt the Shapley-value approach, in which the marginal contribution of a given variable is computed over all the possible sequences of change (Okamoto, 2011; Shorrocks, 2013). An early empirical study based on this approach includes Wan (2004), who analyzed the inequality in China.

The current study differs from most of the existing literature at least in three ways. First, we are considering the *changes* in inequality instead of the level of inequality. When the level of inequality is decomposed, the outcome, which is typically income or consumption, is written as a sums of different "sources", which could be just the product between a covariate and its regression coefficient. Then, inequality is decomposed by sources. The current approach is different because the coefficient is allowed to vary over time, such that we can decompose the change in inequality into the structural component (due to the change in coefficient) and the distributional component (due to the changes in the distribution of covariates). Given that we have the structural and distributional components, the current study is somewhat similar to Fujii (2018), but Fujii (2018) is focused on the ethnic gap, rather than inequality.[7] Second, we explicitly include potentially heterogeneous treatment effect and mediation effect terms. To our knowledge, this is the first study to conduct a regression-based inequality decomposition analysis in an experimental setting. Finally, unlike most studies, we use study hours as an outcome of interest. While average hours of study is surely of interest, inequality measures are also of interest as we would be interested in removing the gap between better and worse students.

We denote the average study hours and indicator for having access to TV or internet ("technology access") by $y_{it}$ and $A_{it}$, respectively, for individual $i$ at time $t \in \{0, 1\}$. We let $X_{it}^j$ be the $j$-th covariate, which may include mediators. In our empirical analysis, we only include the awareness of either online or offline educational programs.[8] We denote the indicator for the treatment group by $T_i$. We then consider the following regression model:

$$y_{it} = \alpha_i + \sum_{a=0}^{1} \mathbf{1}_{(A_{i0}=a)} \cdot \left(\beta_a \mathbf{1}_{(t=1)} + \gamma_a T_i \mathbf{1}_{(t=1)}\right) + \sum_{j=1}^{J} \sum_{\tau=0}^{1} \delta_\tau^j X_{i\tau}^j \mathbf{1}_{(t=\tau)} + \varepsilon_{it}, \qquad (4)$$

[7]Another important difference is that Fujii (2018) uses an integration-based approach rather than the Shapley value for decomposition. The integration-based decomposition has some merits (Fujii, 2017) and can be potentially applied to the current context as well. However, because some of our inequality measures are not differentiable, we chose to use the Shapley decomposition, as it does not require differentiability.

[8]Hours or minutes of educational program viewership were difficult to capture objectively and self-reported viewership information is noisy.

where $\alpha_i$, and $\epsilon_{it}$ represent the individual fixed-effect and idiosyncratic error terms, and $\beta_a$ and $\gamma_a$ capture the time fixed effect and treatment effect, which are allowed to be heterogeneous with respect to the baseline technology access $A_{i0}$. We obtain the estimates of $\alpha_i$, $\beta_a$, $\gamma_a$, $\delta_t^j$ for each $i$, $a$, $j$, and $t$, which will be denoted with a hat ($\hat{}$) by ordinary least-squares regression, and let $\hat{\epsilon}_{it}$ be the regression residual. Now let us define $\hat{S}_{it}^a \equiv \hat{\beta}_a \mathbf{1}_{(t=1)} \mathbf{1}_{(A_{i0}=a)}$ and $\hat{\Gamma}_{it}^a \equiv \hat{\Gamma}_{it}^a \mathbf{1}_{(t=1)} \mathbf{1}_{(A_{i0}=a)}$ for $a \in \{0,1\}$, which can be interpreted as the estimated time-fixed and treatment effects, respectively, for individual $i$ at time $t$ whose status of access to technology at the baseline is $a$. Note that $\hat{S}_{it}^0$ and $\hat{\Gamma}_{it}^0$ [$\hat{S}_{it}^1$ and $\hat{\Gamma}_{it}^1$] are defined to be zero for individuals that have [do not have] access to technology at the baseline (i.e., $A_{i0} = 1$ [$A_{i0} = 0$]). With these notations, we obtain the following identity:

$$
\begin{aligned}
y_{it} &= \hat{\alpha}_i + \sum_{a=0}^{1} \mathbf{1}_{(A_{i0}=a)} \cdot \left( \hat{\beta}_a \mathbf{1}_{(t=1)} + \hat{\gamma}_a T_i \mathbf{1}_{(t=1)} \right) + \sum_{j=1}^{J} \hat{\delta}_t^j X_{it}^j + \hat{\varepsilon}_t \\
&= \hat{\alpha}_i + \hat{S}_{it}^0 + \hat{S}_{it}^1 + \hat{\Gamma}_{it}^0 + \hat{\Gamma}_{it}^1 + \sum_{j=1}^{J} \hat{\delta}_t^j X_{it}^j + \hat{\varepsilon}_t. \qquad (5)
\end{aligned}
$$

In this equation, there is a total of $2J + 5$ variables that have subscript $t$, including $\hat{S}_{it}^0$, $\hat{S}_{it}^1$, $\hat{\Gamma}_{it}^0$, $\hat{\Gamma}_{it}^1$, $\hat{\epsilon}_{it}$, $\hat{\delta}_t^j$, and $X_{it}^j$. Note that we have $y_{i0}$ [$y_{i1}$] when the subscript $t$ for each of these variables is equal to zero [one]. It is also possible to consider $\tilde{y}_i^B$, where some of the variables subscripts $t$ are one and others zero, where $B$ is a $(2J+5)$-digit binary index that represents the combinations of the $t$-subscripts for these $2J + 5$ variables. Table 3 shows four examples. In Row $(a)$ [$(d)$], all subscripts are zero [one]. Hence, $\tilde{y}_i^B$ is equal to $y_{i0}$ [$y_{i1}$]. In Row $(b)$ [$(c)$], the subscript for $\hat{\delta}_t$ is changed from 0 to 1 [1 to 0] in Row (a) [(d)]. Hence, the marginal contribution of $\hat{\delta}_t$ to the change in $y_{it}$ (i.e., $y_{i1} - y_{i0}$) when all other variables' subscript $t$ is equal to zero [one] can be written as $\tilde{y}_i^{(b)} - \tilde{y}_i^{(a)} = \Delta\hat{\delta} X_{i0}$ [ $\tilde{y}_i^{(d)} - \tilde{y}_i^{(c)} = \Delta\hat{\delta} X_{i1}$], where $y^{(m)}$ for $m \in \{a, b, c, d\}$ is the value of $B$ in Row $(m)$ and $\Delta\hat{\delta} \equiv \hat{\delta}_1 - \hat{\delta}_0$ is the change in $\hat{\beta}_t$

While we considered the decomposition of the change in $y_{it}$ here, we can also consider a similar decomposition for the change in the inequality index $I_t \equiv I(\{y_{it}\})$, where $I(\cdot)$ gives an inequality index from the distribution of $y_{it}$ such as the Gini index, general entropy measure, and percentile ratios. As $y_{it}$ can be zero, we only consider measures that allow for the presence of zeros. Therefore, we primarily consider the Gini index ("Gini") as it is a widely used measure of inequality. We also consider the general entropy measure with parameter value of two ("GE2"). We choose the parameter value of two to allow for zero or even negative value in the argument. In addition, we also use the the interquartile ratio ("IQR") and interdecile ratio ("IDR"), which are both percentile ratios. The former [latter] takes the ratio of the 75th [90th] percentile to 25th [10th] percentile. The method discussed below is generally applicable to other well-defined inequality measures.

Table 3: Examples of $B$ and their interpretations

| | Value of $B$ | $q(B)$ | Digit for... $\hat{S}^0_{it}$ | $\hat{S}^1_{it}$ | $\hat{\Gamma}^0_{it}$ | $\hat{\Gamma}^1_{it}$ | $\hat{\epsilon}_{it}$ | $\hat{\delta}^1_t \cdots \hat{\delta}^J_t$ | $X^1_{it} \cdots X^J_{it}$ | $\tilde{y}^B_i$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $(a)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $0\cdots0$ | $0\cdots0$ | $y_{i0}$ |
| $(b)$ | $2^{J+1}-1$ | 1 | 0 | 0 | 0 | 0 | 0 | $1\cdots0$ | $0\cdots0$ | $y_{i0}+\Delta\hat{\delta}^1 X^1_{i0}$ |
| $(c)$ | $2^{2J+5}-2^{J+1}-1$ | 2J+4 | 1 | 1 | 1 | 1 | 1 | $0\cdots1$ | $1\cdots1$ | $y_{i1}-\Delta\hat{\delta}^1 X_{i1}$ |
| $(d)$ | $2^{2J+5}-1$ | 2J+5 | 1 | 1 | 1 | 1 | 1 | $1\cdots1$ | $1\cdots1$ | $y_{i1}$ |

The inequality measures we adopt have different properties. Gini coefficient is bounded between zero and one (in the absence of negative outcomes) and has a straightforward graphical interpretation that it is twice the area between the Lorenz curve and 45-degree line. GE2 is known to be a monotonic transformation of the coefficient of variation and sensitive to the presence of large extreme values. IQR and IDR are robust to the presence of outliers. However, as our method is based on the predictions from linear regression, IDR does not work as well as IQR since the linear models do not predict the tails of the distribution very well. Also, the decomposition would not work (well) if we have a value of (near) zero for some $\tilde{y}$.[9]

To define the Shapley decomposition in our context, we define $\tilde{I}^B \equiv I(\{\tilde{y}^B_i\})$.[10] Now, we consider the marginal effects of each variable by sequentially changing the subscript $t$ and then average the marginal effects across all of the $(2J+5)!$ possible sequences. However, because there are only $2^{2J+5}$ possible combinations of the subscripts, the same marginal effects appears multiple times in this calculation.

To advance the argument, we define $B_{b,l,n}$ be the binary expression where we insert $n \in \{0,1\}$ before the $l$th digit from the left, where $b$ is a $(2J+4)$-digit binary number. Here, $b$ can be interpreted as a summary index that represents the values of subscript $t$ for all variables except for the $l$th variable, for which we are trying to find the marginal contribution. Then, the marginal contribution of the $l$th variable can be written as: $\Delta^l_b \equiv \tilde{I}^{B_{b,l,1}} - \tilde{I}^{B_{b,l,0}}$. Now, let $q(b)$ be the number of digits that take unity in the binary form. Then, there are $q(b)!(2J+4-q(b))!$ sequences along which $\Delta^l_b$ is the relevant marginal effect for the $l$th variable. Therefore, by taking the average over all sequences, we have the marginal contribution of $l$th variable as follows:

[9]This is not an issue for the main analysis using the follow-up survey. However, this is an issue when we use the endline survey for $t=1$.

[10]Note that $\tilde{y}^B_i$ can be negative or more than 24 hours for some $i$ and $B$. Our inequality measures can be defined in the presence of negative values of $\tilde{y}$. Nevertheless, as a robustness check, we define $\check{y} \equiv \min(24, \max(\tilde{y}_{it}, 0))$—a doubly-censored version of $\tilde{y}_{it}$ to be between zero and 24—and use $\check{I}^B \equiv I(\{\check{y}^B_i\})$ for the decomposition analysis. The results remain the same for IQR and IDR and very similar for Gini and GE2.

$$
\begin{aligned}
\Delta^l &= \frac{1}{(2J+5)!}\sum_{b=0}^{2^{2J+4}-1} q(b)!(2J+4-q(b))!\Delta_b^l \\
&= \frac{1}{(2J+5)!}\sum_{B=0}^{2^{2J+5}-1} (q(B)-1)!(2J+4-(q(B)-1))! \cdot \tilde{I}^B \cdot \mathbf{1}_{d(B,l)=1} \\
&\qquad -q(B)!(2J+4-q(B))! \cdot \tilde{I}^B \cdot \mathbf{1}_{d(B,l)=0} \\
&= \sum_{B=0}^{2^{2J+5}-1} \left[\frac{\mathbf{1}_{d(B,l)=1}}{{}_{2J+4}\mathrm{C}_{q(B)-1}} - \frac{\mathbf{1}_{d(B,l)=0}}{{}_{2J+4}\mathrm{C}_{q(B)}}\right] \cdot \frac{\tilde{I}^B}{2J+5}, \qquad (6)
\end{aligned}
$$

where ${}_n\mathrm{C}_k = n!/k!(n-k)!$ is the $n$-choose-$k$ binomial coefficient and $d(B,l)$ is the $l$th digit of $B$ in the binary expression. Because we have $\sum_{l=1}^{2J+5}\Delta_b^l = \Delta_I (\equiv I_1 - I_0)$ for each $b$ by construction, we have $\sum_{l=1}^{2J+5}\Delta^l = \Delta_I$. Eq. (6) is convenient as each decomposed component $\Delta^l$ can be expressed as a weighted sum of the inequality measure for different values of $B$.

Table 4 shows the decomposition results for the change in inequality between the baseline and follow-up.[11] Rows (a) and (b) give the initial and terminal inequality and Row (c) gives the difference between these rows. The decomposition exercise proposed above allows us to account for this difference with different components. Rows (d), (e), (g), (h), (j), (k), and (m) provide different components from the decomposition results, and the percentages in the parentheses represent the share of each decomposition component in the total observed change in Row (c). While the share can be negative or above 100 percent, the sum of $\Delta^1, \cdots \Delta^7$ gives $\Delta_I$ (and thus the shares add up to 100 percent). Table 4 also provides the aggregate components due to time effect (Row (f)), direct treatment effect (Row (i)), and awareness (Row (l)).

The table shows that the inequality in average daily study hours in our sample declined between the baseline and follow-up surveys. As the idiosyncratic effect component is most uninformative, we would like this component to be small. This is the case for Gini, GE2, and IQR, as Row (m) shows. However, for IDR, it is large. As discussed above, this is not surprising as the linear regression does not predict the tails of distribution very well.

With this caveat in mind, let us look at the remaining rows. As Rows (d)–(f) show, much of the decline in inequality is driven by the time effect, particularly for households with technology access. Nevertheless, the treatment also played a role. About 8.50 percent of the decline in Gini can be attributed to the direct effect of treatment (Row (i)). While the share of this component varies across different inequality measures, the total direct treatment effect appears to have contributed to 5-20 percent of the decline in

[11] The change in inequality between the baseline and endline is generally small and mixed. Gini and IDR increased by 0.0084 and 0.500, respectively, whereas GE2 and IQR decreased by 0.0164 and 0.6667, respectively.

Table 4: Decomposition results

| Inequality index | | Gini | GE2 | IQR | IDR |
|---|---|---|---|---|---|
| (a) | Initial inequality ($I_0$) | 0.2867 | 0.1546 | 3.0000 | 4.0000 |
| (b) | Terminal inequality ($I_1$) | 0.1996 | 0.0628 | 1.7368 | 2.7857 |
| (c) | Change in inequality ($\Delta_I$) | -0.0871 | -0.0918 | -1.2632 | -1.2143 |
| (d) | Time effect for HH | -0.0196 | -0.0172 | -0.3008 | -0.1570 |
| | without TV/inet access ($\Delta^1$) | (22.52%) | (18.71%) | (23.81%) | (12.93%) |
| (e) | Time effect for HH | -0.0581 | -0.0485 | -0.3880 | -1.5464 |
| | with TV/inet access ($\Delta^2$) | (66.67%) | (52.84%) | (30.72%) | (127.35%) |
| (f) | Direct treatment effect on | -0.0031 | -0.0037 | -0.1683 | 0.0055 |
| | HH without TV/inet access ($\Delta^3$) | (3.52%) | (3.98%) | (13.33%) | (-0.46%) |
| (g) | Direct treatment effect on | -0.0043 | -0.0048 | -0.0917 | -0.0733 |
| | HH with TV/inet access ($\Delta^4$) | (4.98%) | (5.28%) | (7.26%) | (6.04%) |
| (h) | Structural effect | -0.0084 | -0.0098 | -0.2293 | 0.0512 |
| | of awareness ($\Delta^5$) | (9.63%) | (10.73%) | (18.15%) | (-4.22%) |
| (i) | Distributional effect | -0.0021 | -0.0026 | -0.0895 | 0.0772 |
| | of awareness ($\Delta^6$) | (2.43%) | (2.84%) | (7.08%) | (-6.36%) |
| (j) | Idiosyncratic | 0.0085 | -0.0052 | 0.0045 | 0.4285 |
| | effect ($\Delta^7$) | (-9.75%) | (5.63%) | (-0.35%) | (-35.29%) |
| (k) | Total time effect | -0.0777 | -0.0657 | -0.6888 | -1.7034 |
| | ($\Delta^1 + \Delta^2$) | (89.19%) | (71.56%) | (54.53%) | (140.28%) |
| (l) | Total direct treatment | -0.0074 | -0.0085 | -0.2601 | -0.0678 |
| | effect ($\Delta^3 + \Delta^4$) | (8.50%) | (9.25%) | (20.59%) | (5.58%) |
| (m) | Total awareness effect | -0.0105 | -0.0125 | -0.3188 | 0.1284 |
| | effect ($\Delta^5 + \Delta^6$) | (12.06%) | (13.57%) | (25.24%) | (-10.58%) |

Tech access refers to the access to either TV or internet at home. The percentages in the parentheses show the size of the inequality component relative to the change ($\Delta_I$). IQR and IDR, respectively, stand for interquartile ratio (p75/p25) and interdecile ratio (p90/p10).

inequality.

The total direct treatment effect is arguably an underestimate of the overall treatment effect because the former does not account for the changes in awareness that might be affected by the treatment and contributed to the study hours. For example, if we believe that the distributional effect of awareness can be attributed to the treatment, 10.93(=8.50+2.43) percent of the decline in Gini could be attributed to the treatment. The results are qualitatively similar for GE2 and IQR as well. However, a similar calculation leads to -0.78(=5.58-6.36) percent for IDR, but because the proportion of idiosyncratic effect for IDR is much larger than other inequality measures (Row (j)), this number is not as reliable as other inequality measures. Taken together, our treatment appears to have reduced the inequality in study hours in the middle of the distribution but did not affect the inequality between the tails.

# 6 Conclusion

The COVID-19 pandemic has negatively affected the lives of children worldwide, especially in Bangladesh. Schools were closed for a year and a half, and even as they gradually reopened and resumed regular schedules, it was challenging for students to fully return to their pre-pandemic routine educational activities. The Government of Bangladesh has been providing educational content through telecasting and online platforms to address the learning loss due to extended school closure. However, the uptake of these programs has been abysmally low, primarily due to the lack of awareness and commitment. Our intervention aims to study the impact of high-frequency information nudges on the uptake of online and offline educational content and study behavior for secondary school students in rural Bangladesh.

Our intervention increases the awareness of the educational content and the average daily study hours, particularly for disadvantaged children. As a result, our intervention also contributes to reducing inequality in study hours. It is scalable as our intervention exploits widespread mobile technology to improve the awareness, viewership, and utilization of readily available educational content. It could be replicated in other contexts with comparable educational content available to the public. Besides the scalability, our results also suggest that our intervention can reduce inequality in the study hours. Based on our decomposition analysis, about 5-20 percent of the reduction in inequality between the baseline and follow-up surveys can be attributed to our intervention's direct and indirect effects.

A few important policy implications can be derived from this study. First, providing educational content open to the public alone is unlikely to help the disadvantaged group by itself. The target group must be aware of the presence of such content to produce a meaningful impact. While this appears to be self-evident, the current study underscores the importance of awareness. Second, it may take some time for the effects of information nudges to become detectable, as human behavior tends to follow the status quo. We only found a significant impact on study hours at the follow-up, but neither the study hours nor the test score increased significantly at the endline. Third, because even the parents of disadvantaged children tend to direct their children to private tutors, the educational content that is currently available would not be sufficient. With the improvement in personalized learning coupled with the spread of smartphones and computers, these students can benefit more from these technologies if made available for free or at a moderate cost. Such an intervention could further decrease the inequality in study behavior and make the playing field level between students with different backgrounds.

# References


Alon, S. (2007). The influence of financial aid in leveling group differences in graduating from elite institutions. *Economics of Education Review*, 26(3):296–311.

Amin, S., Hossain, M., and Ainull, S. (2021). Learning loss among adolescent girls during the COVID-19 pandemic in rural Bangladesh. Brief, The Population Council. `https://www.popcouncil.org/uploads/pdfs/2021SBSR_LearningLossBangladesh.pdf` (Accessed on May 12, 2022).

Angrist, N., Bergman, P., Brewster, C., and Matsheng, M. (2020). Stemming learning loss during the pandemic: A rapid randomized trial of a low-tech intervention in Botswana. Available at SSRN: `https://ssrn.com/abstract=3663098`.

Asadullah, M. N., Trannoy, A., Tubeuf, S., and Yalonetzky, G. (2021). Measuring educational inequality of opportunity: Pupil's effort matters. *World Development*, 138:105262.

Asadullah, N., Bhattacharjee, A., Tasnim, M., and Mumtahena, F. (2020). COVID-19, schooling and learning. BRAC Institute of Governance and Development. `https://bigd.bracu.ac.bd/wpcontent/uploads/2020/06/COVID-19~?` (Accessed on December 12, 2021).

Avvisati, F., Gurgand, M., Guyon, N., and Maurin, E. (2014). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies*, 81(1):57–83.

Baird, S., Seager, J., Sabarwal, S., Guglielmi, S., and Sultan, M. (2020). Adolescence in the time of COVID-19: Evidence from Bangladesh. Policy brief: November 2020, South Asia Gender Innovation Lab. `https://www.gage.odi.org/publication/adolescence-in-the-time-of-covid-19-evidence-from-bangladesh/`.

Barrera-Osorio, F., Gertler, P., Nakajima, N., and Patrinos, H. (2020a). Promoting parental involvement in schools: Evidence from two randomized experiments. NBER Working Paper 28040, National Bureau of Economic Research.

Barrera-Osorio, F., Gonzalez, K., Lagos, F., and Deming, D. J. (2020b). Providing performance information in education: An experimental evaluation in Colombia. *Journal of Public Economics*, 186:104185.

Bergman, P. (2021). Parent-child information frictions and human capital investment: Evidence from a field experiment. *Journal of Political Economy*, 129(1):286–322.

Bergman, P. and 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., and Martinez, C. (2021). Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile. NBER Working Paper 28581, National Bureau of Economic Research.

Betthäuser, B. A., Bach-Mortensen, A. M., and Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, 7(3):375–385.

Bhuiyan, Mohsin (2021). 62 percent Bangladeshi users to have smartphones by 2025: Report. The Business Standard. `https://www.tbsnews.net/tech/62-bangladeshi-users-have-smartphones-2025-report-294121` (Accessed on April 7, 2022).

Biswas, K., Asaduzzaman, T., Evans, D. K., Fehrler, S., Ramachandran, D., and Sabarwal, S. (2020). *TV-Based Learning in Bangladesh: Is it Reaching Students?* World Bank, Washington, DC.

Blinder, A. (1973). Age discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8:436–455.

Brisson, B. M., Dicke, A.-L., Gaspard, H., Häfner, I., Flunger, B., Nagengast, B., and Trautwein, U. (2017). Short intervention, sustained effects: Promoting students' math competence beliefs, effort, and achievement. *American Educational Research Journal*, 54(6):1048–1078.

Castleman, B. L. and Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115:144–160.

Cheney, Catherine (2018). Want to reach the world's poorest? Design for dumb phones. Devex. `https://www.devex.com/news/want-to-reach-the-world-s-poorest-design-for-dumb-phones-90993` (Accessed on April 7, 2022).

Clark, A. E., Nong, H., Zhu, H., and Zhu, R. (2021). Compensating for academic loss: Online learning and student performance during the COVID-19 pandemic. *China Economic Review*, 68:101629.

Datta, S. K. (2022). Online education in Bangladesh during COVID-19: Problems and possibilities. In *Handbook of Research on Asian Perspectives of the Educational Impact of COVID-19*, pages 223–241. IGI Global.

De Walque, D. and Valente, C. (2018). Incentivizing school attendance in the presence of parent-child information frictions. World Bank Policy Research Working Paper 8476, World Bank.

Di Pietro, G., Biagi, F., Costa, P., Karpiński, Z., and Mazza, J. (2020). *The likely impact of COVID-19 on education: Reflections based on the existing literature and recent international datasets*, volume 30275. Publications Office of the European Union Luxembourg.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, 109(8):2728–65.

Donnelly, R. and Patrinos, H. A. (2021). Learning loss during COVID-19: An early systematic review. *Prospects*, pages 1–9.

Engzell, P., Frey, A., and Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(17).

Fack, G. and Grenet, J. (2015). Improving college access and success for low-income students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics*, 7(2):1–34.

Fujii, T. (2017). Dynamic poverty decomposition analysis: An application to the Philippines. *World Development*, 100:69–84.

Fujii, T. (2018). Has the development gap between the ethnic minority and majority groups narrowed in Vietnam?: Evidence from household surveys. *World Economy*, 41(8):2067–2101.

Fujii, T., Ho, C., Ray, R., and Shonchoy, A. (2023). School attendance information or conditional cash transfer? Evidence from a randomized field experiment in rural Bangladesh. Research Collection School Of Economics, 1-2023.

Hanushek, E. A. and Woessmann, L. (2020). *The economic impacts of learning losses*. OECD. https://www.oecd.org/education/The-economic-impacts-of-coronavirus-covid-19-learning-losses.pdf (Accessed on May 1, 2022).

Hassan, M. H., Islam, A., Siddique, A., Wang, L. C., et al. (2023). *Telementoring and homeschooling during school closures: A randomized experiment in rural Bangladesh*. IZA-Institute of Labor Economics.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. routledge.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3):410–442.

Khan, M. M., Rahman, S. T., and Islam, S. T. A. (2021). Online education system in Bangladesh during COVID-19 pandemic. *Creative Education*, 12(2):441–452.

Kraft, M. A. and Dougherty, S. M. (2013). The effect of teacher–family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness*, 6(3):199–222.

Kraft, M. A. and Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review*, 47:49–63.

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

Mahmud, F. (2021). Bangladesh reopens schools after 18-month COVID shutdown. Al Jazeera. `https://www.aljazeera.com/news/2021/9/13/bangladesh-reopens-schools-after-18-month-covid-shutdown`.

Makino, M., Shonchoy, A., and Wahhaj, Z. (2021). Early effects of the COVID-19 lockdown on children in rural Bangladesh. Studies in Economics 2102, School of Economics, University of Kent.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14:693–709.

OECD (2021). The state of school education: One year into the COVID pandemic.

Okamoto, M. (2011). Source decomposition of changes in income inequality: The integral-based approach and its approximation by the chained Shapley-value approach. *Journal of Economic Inequality*, 9:145–181.

Rogers, T. and Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour*, 2(5):335–342.

Sarwar, A., Hossain, I., and Kaye, T. (2020). Five lessons learnt from Bangladesh's experience responding to COVID-19. EdTech Hub. `https://edtechhub.org/2020/07/15/five-lessons-learnt-from-bangladeshs-experience-responding-to-covid-19s/` (Accessed on May 17, 2022).

Shorrocks, A. (2013). Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. *Journal of Economic Inequality*, 11:99–126.

Silver, L. and Johnson, C. (2018). Majorities in Sub-Saharan Africa own mobile phones, but smartphone adoption is modest. Pew Research Centre: Global attitudes & trends. `https://www.pewresearch.org/global/2018/10/09/majorities-in-sub-saharan-africa-own-mobile-phones-but-smartphone/` (Accessed on May 1, 2022).

UNESCO (2016). *Global Education Monitoring Report 2016: Education for People and Planet-Creating Sustainable Futures for All.* UN.

UNICEF (2020). Students in Bangladesh adjust to remote learning via national TV during COVID-19 lockdown. `https://www.unicef.org/bangladesh/en/stories/` (Accessed on May 1, 2022).

UNICEF et al. (2021). COVID-19 and school closures. *One year of education disruption. UNICEF.*

Wan, G. (2004). Accounting for income inequality in rural China: A regression-based approach. *Journal of Comparative Economics*, 32:348–363.

World Bank (2016). *World Development Report 2016: Digital Dividends.* World Bank.

World Bank (2020). Poverty maps (Bangladesh - Admin 2 and 3 - 2010). Last Updated on August 3, 2021, 1:23AM. `https://data.apps.fao.org/catalog/dataset/bangladesh-interactive-poverty-maps` (Accessed on December 12, 2021).

York, B. N., Loeb, S., and 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.

Yun, M.-S. (2006). Earnings inequality in USA, 1969–99: Comparing inequality using earnings equations. *Review of Income and Wealth*, 52(1):127–144.

# A Appendix

## A.1 Power Calculation

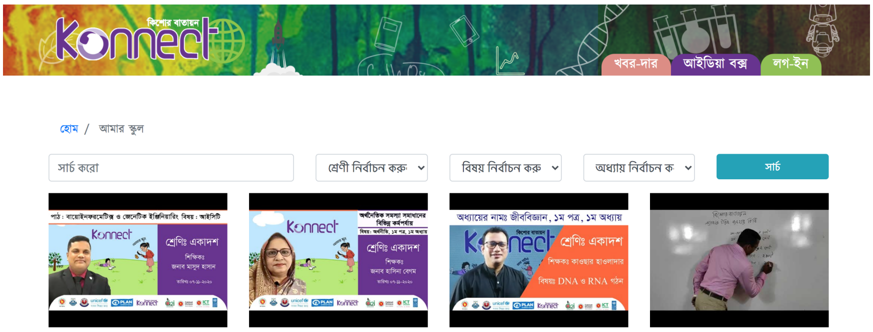
Since our intervention is highly unlikely to decrease study hours, we conducted a one-sided test. We denote the significance level and the probability of type-II error by $\alpha$ and $\beta$, respectively. Let $P$ and $N$ respectively denote the proportion of treated individuals and the sample size. Then, because our study design was based on individual level randomization, we applied the following standard formula under normal approximation to calculate the minimum detectable effect size (MDES)—expressed in the multiples of the population standard deviation of the outcome of interest:

$$\text{MDES} = \frac{z_{1-\alpha} + z_{1-\beta}}{\sqrt{NP(1-P)}}, \tag{A1}$$

where $z_p$ is the $z$-critical value for probability $p$ for standard normal distribution. Since we recruited a total of 1,200 students and divided them equally into the information treatment and control arms, we set $N = 1,200$ and $P = 0.5$. Further, letting $\alpha = 0.1$, $1-\beta = 0.80$, we have MDES $\simeq 0.12$ standard deviation. Therefore, if we find statistically significant impact, it will likely be economically significant as well.

## A.2 Additional Figures and Tables

Figure A1: Examples of Online and Offline Educational Content




Panel A: Konnect

গণপ্রজাতন্ত্রী বাংলাদেশ সরকার
মাধ্যমিক ও উচ্চ শিক্ষা অধিদপ্তর
বাংলাদেশ, ঢাকা
www.dshe.gov.bd

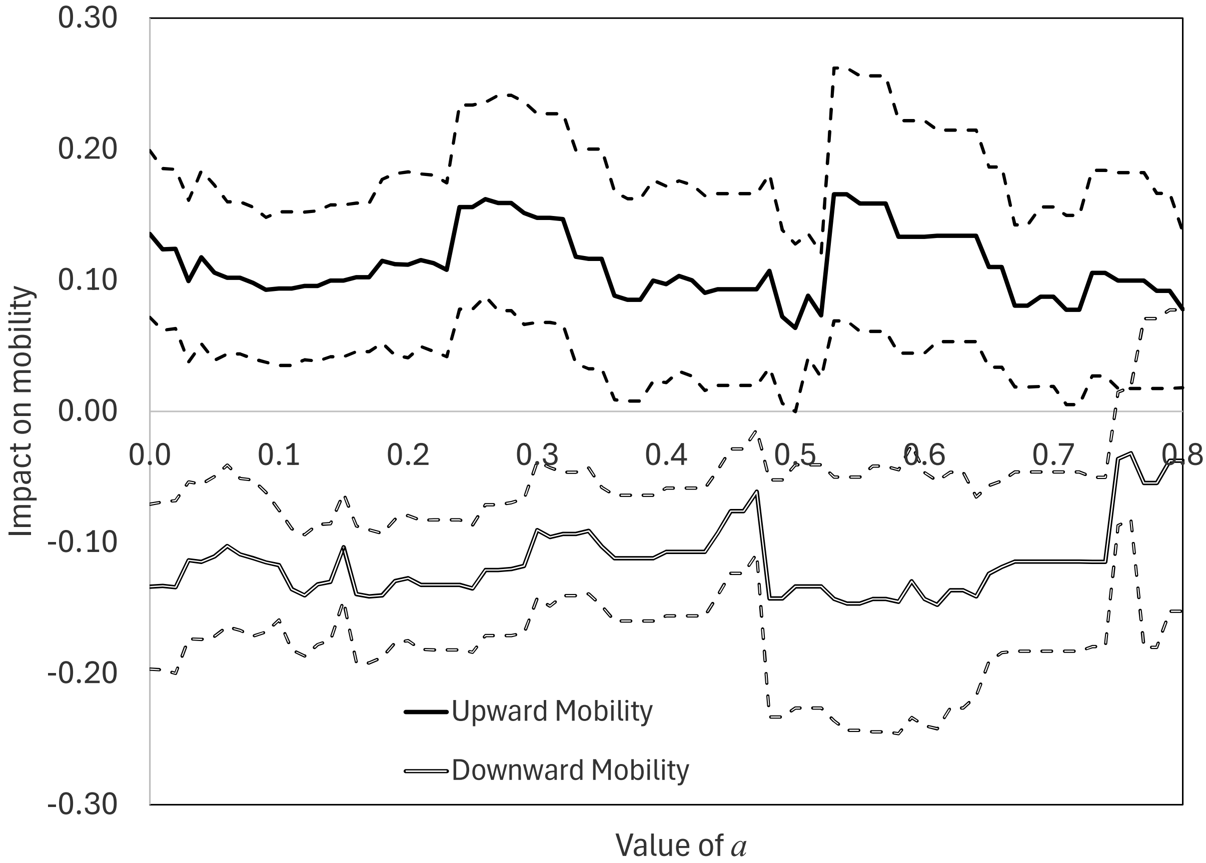
**ক্লাস রুটিন**
**'আমার ঘরে আমার স্কুল'**

| তারিখ | শ্রেণি | ১ম পিরিয়ড | ২য় পিরিয়ড |
|---|---|---|---|
| ১৯ এপ্রিল-২০২০ | জাতীয় সঙ্গীত ও করোনা সচেতনতা (সকাল ১১.০০-১১.০৫) | | |
| | ষষ্ঠ (সকাল ১১.০৫-১১.৪৫) | বিজ্ঞান | বাংলা |
| | সপ্তম (সকাল ১১.৪৫-দুপুর ১২.২৫) | গণিত | বিজ্ঞান |
| | অষ্টম (দুপুর ১২.২৫-১.০৫) | বাংলা | বিজ্ঞান |
| | নবম (দুপুর ১.০৫- ১.৪৫) | গণিত | পদার্থ বিজ্ঞান |
| | দশম (দুপুর ১.৪৫- ২.২৫) | ইংরেজি | রসায়ন বিজ্ঞান |
| ২০ এপ্রিল-২০২০ | জাতীয় সঙ্গীত ও করোনা সচেতনতা (সকাল ১১.০০-১১.০৫) | | |
| | ষষ্ঠ (সকাল ১১.০৫-১১.৪৫) | গণিত | বাংলাদেশ ও বিশ্বপরিচয় |
| | সপ্তম (সকাল ১১.৪৫-দুপুর ১২.২৫) | ইংরেজি | গণিত |
| | অষ্টম (দুপুর ১২.২৫-১.০৫) | বিজ্ঞান | বাংলাদেশ ও বিশ্বপরিচয় |
| | নবম (দুপুর ১.০৫- ১.৪৫) | জীব বিজ্ঞান | ইংরেজি |
| | দশম (দুপুর ১.৪৫- ২.২৫) | গণিত | ভূগোল |
| ২১ এপ্রিল-২০২০ | জাতীয় সঙ্গীত ও করোনা সচেতনতা (সকাল ১১.০০-১১.০৫) | | |
| | ষষ্ঠ (সকাল ১১.০৫-১১.৪৫) | বাংলা | গণিত |
| | সপ্তম (সকাল ১১.৪৫-দুপুর ১২.২৫) | গণিত | বাংলাদেশ ও বিশ্বপরিচয় |
| | অষ্টম (দুপুর ১২.২৫-১.০৫) | বাংলা | গণিত |
| | নবম (দুপুর ১.০৫- ১.৪৫) | গণিত | ইংরেজি |
| | দশম (দুপুর ১.৪৫- ২.২৫) | ইংরেজি | রসায়ন |

Panel B: Sangsad TV Schedule

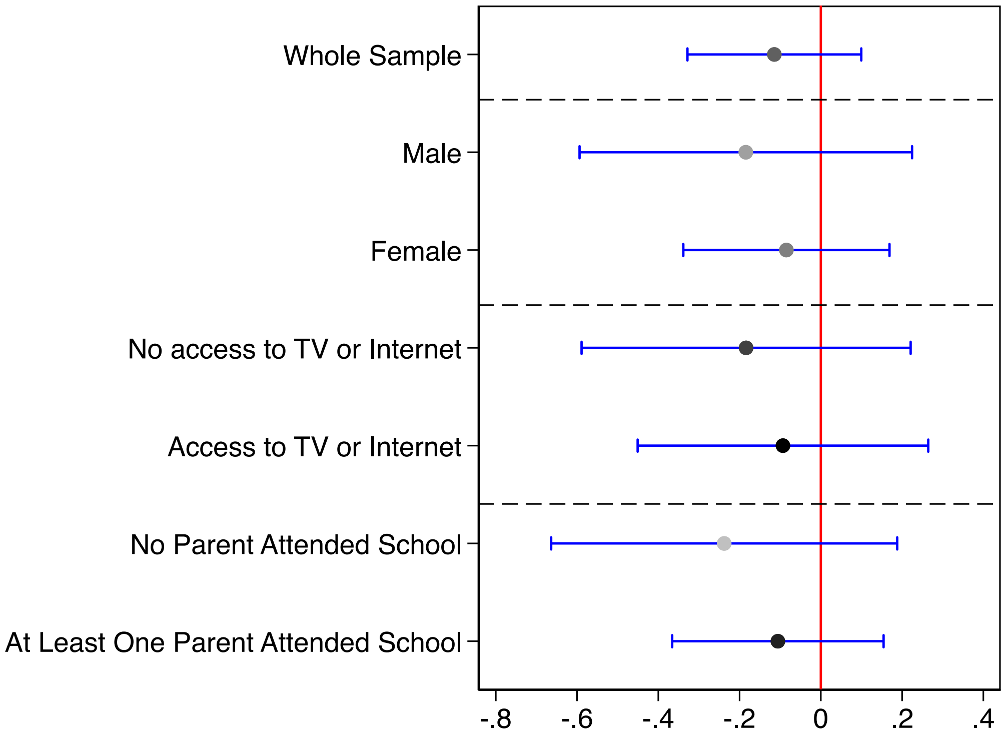
Note: Panel A is an illustration of the home page for the Konnect website (http://konnect.edu.bd/my-school) and Panel B shows the subject-grade specific schedule of educational programs on Sangsad TV for April, 2020, as published by the Ministry of Education in Bangladesh.

Figure A2: Treatment effect on upward and downward mobility in daily study hours for different values of $a$



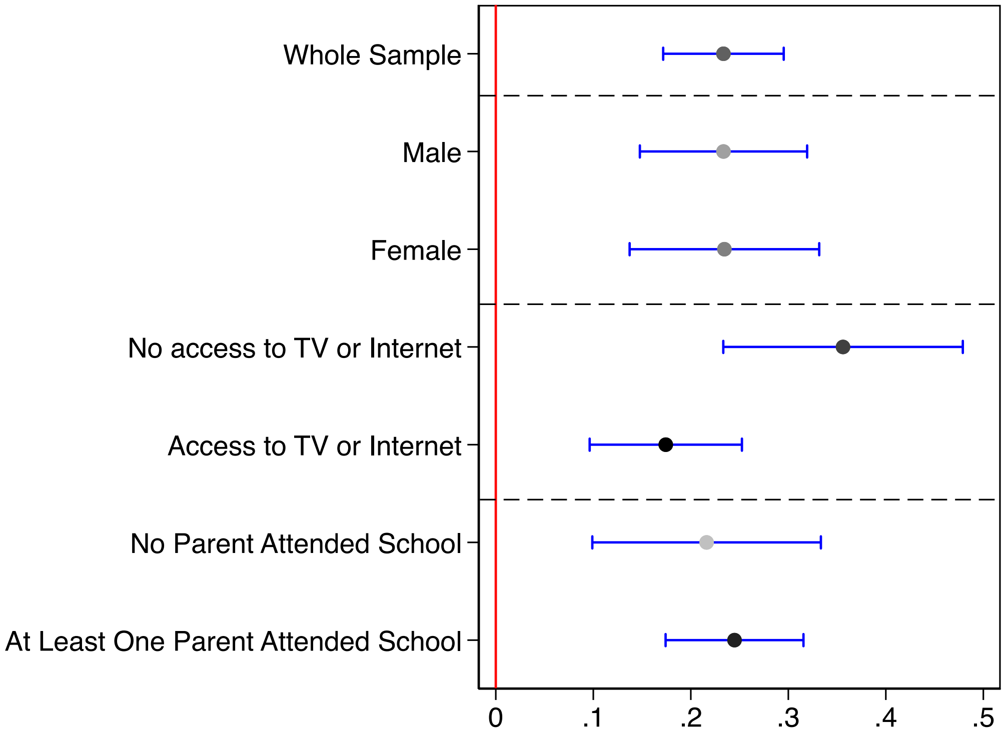

Note: The number of observations varies with the value of $a$, because the observations that are within the top [bottom] $100a$ percentile in the control distribution of average daily study hours at the baseline are dropped from the analysis. The upward and downward mobility measures are defined in eq. (2). The unbalanced covariate at the baseline–ownership of bicycle–and school-grade-gender fixed-effects are controlled for. The dashed lines represent 95-percent confidence interval based on robust standard errors clustered at the school-grade-gender level.

Figure A3: Impact of Treatment on Test Scores



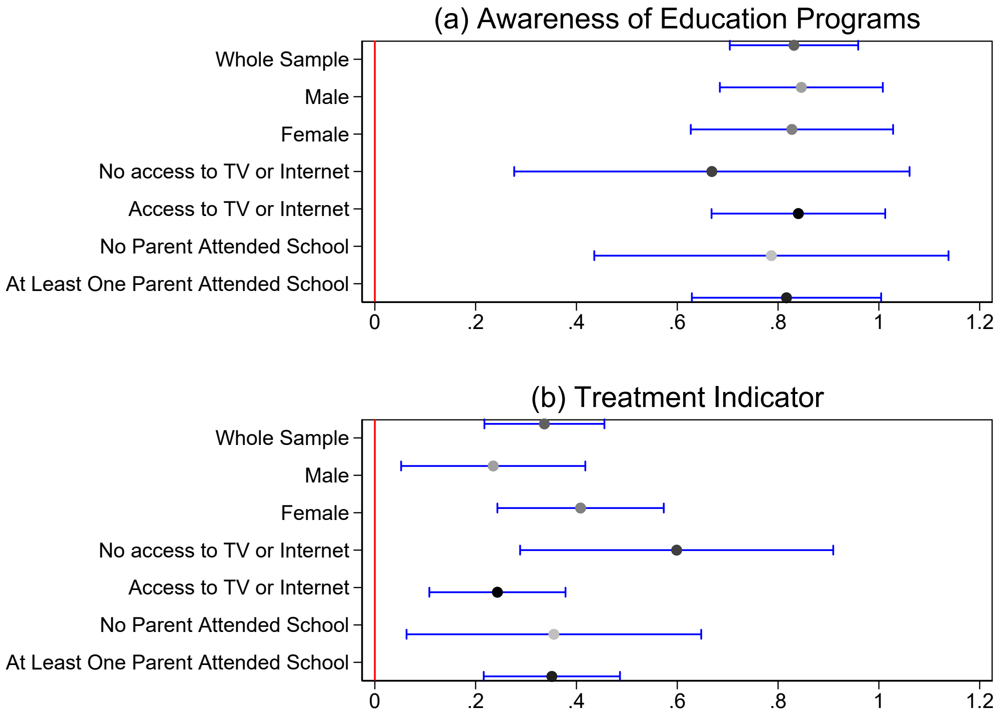

Note: The above estimates are derived from the endline survey conducted with 1,137 students. The outcome variable is the test score at the endline. The specifications control for test score at the baseline, unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

Figure A4: Impact of Treatment on Awareness About Education Programs




Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable in the above figure is awareness of education programs on online or offline platforms. It takes unity if the parent is aware of education programs on offline platforms such as Sangsad TV or online programs such as Youtube, Facebook, Konnect or Whatsapp, and 0 otherwise. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

Figure A5: Mediation Analysis




Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. Figures A5(a) and (b) plot the coefficients and confidence intervals on the indicators for the awareness of education programs and treatment, respectively, in the mediation regression analysis. The outcome variable in the above figures is the average daily study hours in the past year as reported by parents in the follow-up survey. The specifications control for the unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. The 95-percent confidence intervals are based on the standard errors clustered at the school-grade-gender level.

Table A1: Examples of High Frequency Information

| Sl | Message |
|---|---|
| 1 | Greetings from MOMODa Foundation! You may consult your English grammar teacher <insert teacher name> over message/voice call at <insert phone number> if you have any English grammar related concern. Keep studying! |
| 2 | Greetings from MOMODa Foundation! You can watch education programs for <insert grade of student> on Sangsad TV on <insert day of the week> at <insert time (in HHMM format)> every week. Keep studying! |
| 3 | Greetings from MOMODa Foundation! Research finds that one additional year of schooling improves wages by 7 percent. Keep studying! |
| 4 | Greetings from MOMODa Foundation! You can access educational content on <insert educational platform> at <insert hyperlink>. Keep studying! |

Table A2: Compliance Rate of Intervention

| **Week** | SMS | Voice Call | **Week** | SMS | Voice Call | **Week** | SMS | Voice Call |
|---|---|---|---|---|---|---|---|---|
| **1** | 100 | 66 | **11** | 100 | 70 | **21** | 100 | 89 |
| **2** | 100 | 75 | **12** | 97 | 17 | **22** | 100 | 88 |
| **3** | 100 | 23 | **13** | 100 | 87 | **23** | 89 | 63 |
| **4** | 100 | 20 | **14** | 100 | 73 | **24** | 89 | 77 |
| **5** | 100 | 91 | **15** | 98 | 87 | **25** | 100 | 86 |
| **6** | 100 | 89 | **16** | 96 | 88 | **26** | 100 | 84 |
| **7** | 100 | 100 | **17** | 100 | 87 | **27** | 100 | 77 |
| **8** | 100 | 100 | **18** | 97 | 89 | **28** | 100 | 75 |
| **9** | 99 | 83 | **19** | 100 | 83 | **29** | 100 | 76 |
| **10** | 64 | 56 | **20** | 78 | 87 | **30** | 100 | 83 |

Note: The above table provides the average compliance rate of the intervention in the treatment group—SMS and voice calls received—by week.

Table A3: Correlations Across Variables of Interest

| | Study Hours Reported by Parent at the baseline | Study Hours Reported by Parent at the endline | Study Hours Reported by Parent at follow-up | Study Hours Reported by Student at the baseline | Study Hours Reported by Student at the endline | Student Motivation at Baseline | Student Motivation at Endline | Test Score at Baseline | Test Score at Endline |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Study Hours Reported by Parent at the baseline | 1.000 | | | | | | | | |
| Study Hours Reported by Parent at the endline | 0.016 | 1.000 | | | | | | | |
| Study Hours Reported by Parent at follow-up | 0.114** | 0.031 | 1.000 | | | | | | |
| Study Hours Reported by Student at the baseline | 0.521*** | -0.005 | 0.107** | 1.000 | | | | | |
| Study Hours Reported by Student at the endline | 0.133*** | 0.439*** | 0.161*** | 0.115** | 1.000 | | | | |
| Student Motivation at Baseline | 0.076* | -0.063 | 0.061 | 0.009 | 0.044 | 1.000 | | | |
| Student Motivation at Endline | 0.008 | 0.068 | 0.068 | 0.057 | -0.017 | 0.024 | 1.000 | | |
| Test Score at Baseline | 0.109** | 0.033 | 0.106** | 0.090** | 0.075* | 0.064 | 0.036 | 1.000 | |
| Test Score at Endline | 0.119*** | 0.042 | 0.172*** | 0.132*** | 0.071* | 0.040 | 0.103** | 0.295*** | 1.000 |

Note: Columns (1)-(5) use daily study hours in different survey rounds. Columns (6) and (7) use the raw score for motivation of students at the baseline and endline, respectively. Motivation scores can range between 14 to 70. Columns (8) and (9) use the raw score of students in the Mathematics test administered by the research team at the baseline, and endline respectively. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A4: Impact of Treatment on Discontinuation

| Dependent variable | Discontinued | |
|---|---|---|
| | Endline | Follow-up |
| | (1) | (2) |
| Treated | -0.013 | -0.021 |
| | (0.014) | (0.017) |
| | [0.511] | [0.278] |
| Student observations | 1,200 | 1,200 |
| $R^2$ | 0.042 | 0.028 |
| Control Mean | 0.058 | 0.112 |

Note: "Discontinued" takes unity if the child discontinued from the study at the endline or follow-up, and 0 otherwise. 62 students discontinued at the endline and 122 students discontinued at follow-up. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A5: Impact of Treatment on Study Hours: ITT and 2SLS Estimates

| Dependent variable | Daily Study Hours in Past One Year | | | | | | |
|---|---|---|---|---|---|---|---|
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **ITT Estimates** | | | | | | | |
| Treated | 0.531*** | 0.433*** | 0.602*** | 0.837*** | 0.390*** | 0.525*** | 0.551*** |
| | (0.057) | (0.076) | (0.079) | (0.133) | (0.066) | (0.129) | (0.069) |
| | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] |
| **2SLS Estimates** | | | | | | | |
| Treated | 0.765*** | 0.649*** | 0.842*** | 1.175 *** | 0.573*** | 0.862*** | 0.763*** |
| | (0.110) | (0.180) | (0.138) | (0.193) | (0.134) | (0.231) | (0.123) |
| | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.002] | [0.001] |
| Student observations | 1,077 | 471 | 606 | 313 | 764 | 272 | 805 |
| Control Mean | 3.525 | 3.562 | 3.497 | 3.267 | 3.631 | 3.418 | 3.560 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable in the above figure is a continuous variable measuring daily study hours in the past one year. For the 2SLS estimates, we instrument for compliance, that is, whether any member in the treated household received SMS or voice call, with treatment assignment. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A6: Impact of Treatment on Mobility

| Dependent variable | Upward Mobility | | | | | | |
|---|---|---|---|---|---|---|---|
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treated | 0.135*** | 0.082* | 0.177*** | 0.175** | 0.118*** | 0.189** | 0.124*** |
| | (0.031) | (0.040) | (0.044) | (0.066) | (0.034) | (0.070) | (0.032) |
| | [0.001] | [0.067] | [0.002] | [0.013] | [0.002] | [0.013] | [0.001] |
| Control Mean | 0.462 | 0.478 | 0.450 | 0.400 | 0.488 | 0.408 | 0.480 |
| Dependent variable | Downward Mobility | | | | | | |
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treated | -0.134*** | -0.078* | -0.177*** | -0.175** | -0.116*** | -0.189 ** | -0.121*** |
| | (0.031) | (0.039) | (0.044) | (0.066) | (0.033) | (0.070) | (0.032) |
| | [0.001] | [0.069] | [0.002] | [0.013] | [0.002] | [0.013] | [0.001] |
| Control Mean | 0.536 | 0.517 | 0.550 | 0.600 | 0.509 | 0.592 | 0.518 |
| Student observations | 1,072 | 469 | 603 | 313 | 759 | 269 | 803 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable in the above figure is a mobility measures based on study hours at the baseline and follow-up. "Upward mobility" ["Downward mobility"] is an indicator variable that takes a value of 1 if the student moved up [down] in the control distribution of average daily study hours between the baseline and follow-up surveys. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A7: Impact of Treatment on Test Scores

| Dependent variable | Test Score at Endline | | | | | | |
|---|---|---|---|---|---|---|---|
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treated | -0.114 | -0.185 | -0.085 | -0.184 | -0.093 | -0.238 | -0.106 |
| | (0.105) | (0.191) | (0.118) | (0.197) | (0.175) | (0.208) | (0.127) |
| | [0.369] | [0.287] | [1.000] | [0.439] | [1.000] | [1.000] | [0.637] |
| Student observations | 1,137 | 486 | 651 | 344 | 793 | 298 | 839 |
| Control Mean | 4.491 | 4.772 | 4.273 | 4.297 | 4.570 | 4.206 | 4.584 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable is the raw test score in the mathematics exam conducted at endline. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A8: Impact of Treatment on Awareness About Education Programs

| Dependent variable | Awareness About Education Programs | | | | | | |
|---|---|---|---|---|---|---|---|
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treated | 0.233*** | 0.233*** | 0.234*** | 0.356*** | 0.174*** | 0.216*** | 0.245*** |
| | (0.030) | (0.040) | (0.045) | (0.060) | (0.038) | (0.057) | (0.035) |
| | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] |
| Student observations | 1,077 | 471 | 606 | 313 | 764 | 272 | 805 |
| Control Mean | 0.382 | 0.392 | 0.373 | 0.271 | 0.427 | 0.344 | 0.394 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable takes unity if the parent is aware of education programs on offline platforms such as Sangsad TV or online programs such as Youtube, Facebook, Konnect or Whatsapp, and 0 otherwise. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A9: Mechanisms: Sources of Help Directed by Parents

| Dependent variable | Education Platforms | | | | | | |
|---|---|---|---|---|---|---|---|
| | All (1) | Male (2) | Female (3) | No Access to TV or Internet (4) | Has Access to TV or Internet (5) | No Parent Attended School (6) | At Least One Parent Attended School (7) |
| Treated | 0.227*** | 0.237*** | 0.220*** | 0.245*** | 0.212*** | 0.161*** | 0.251*** |
| | (0.025) | (0.037) | (0.034) | (0.068) | (0.032) | (0.053) | (0.031) |
| | [0.001] | [0.001] | [0.001] | [0.002] | [0.001] | [0.006] | [0.001] |
| Control Mean | 0.130 | 0.138 | 0.123 | 0.097 | 0.143 | 0.137 | 0.127 |
| Dependent variable | School Teachers | | | | | | |
| | All (1) | Male (2) | Female (3) | No Access to TV or Internet (4) | Has Access to TV or Internet (5) | No Parent Attended School (6) | At Least One Parent Attended School (7) |
| Treated | -0.009 | -0.005 | -0.011 | -0.072 | 0.013 | -0.063 | 0.010 |
| | (0.035) | (0.073) | (0.028) | (0.064) | (0.040) | (0.081) | (0.035) |
| | [1.000] | [1.000] | [1.000] | [0.365] | [1.000] | [0.787] | [1.000] |
| Control Mean | 0.707 | 0.728 | 0.690 | 0.781 | 0.676 | 0.725 | 0.701 |
| Dependent variable | Private Tutoring | | | | | | |
| | All (1) | Male (2) | Female (3) | No Access to TV or Internet (4) | Has Access to TV or Internet (5) | No Parent Attended School (6) | At Least One Parent Attended School (7) |
| Treated | 0.071*** | 0.045 | 0.093** | 0.094* | 0.053* | 0.027 | 0.087*** |
| | (0.022) | (0.027) | (0.034) | (0.048) | (0.028) | (0.054) | (0.020) |
| | [0.004] | [0.141] | [0.016] | [0.063] | [0.076] | [1.000] | [0.001] |
| Control Mean | 0.761 | 0.789 | 0.740 | 0.716 | 0.780 | 0.786 | 0.753 |
| Student observations | 1,077 | 471 | 606 | 313 | 764 | 272 | 805 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variables in the above panel take unity if the parents guide their children to the mentioned source for study-related help in the past year, and zero otherwise. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.

Table A10: Mediation Analysis

| Dependent variable | Awareness About Education Programs | | | | | | |
|---|---|---|---|---|---|---|---|
| | All | Male | Female | No Access to TV or Internet | Has Access to TV or Internet | No Parent Attended School | At Least One Parent Attended School |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treated | 0.336*** | 0.235** | 0.408*** | 0.599*** | 0.243*** | 0.355** | 0.351*** |
| | (0.058) | (0.085) | (0.077) | (0.152) | (0.066) | (0.142) | (0.066) |
| | [0.001] | [0.008] | [0.001] | [0.001] | [0.001] | [0.010] | [0.001] |
| Awareness of Programs | 0.832*** | 0.846*** | 0.828*** | 0.669*** | 0.840*** | 0.787*** | 0.817*** |
| | (0.062) | (0.075) | (0.094) | (0.191) | (0.084) | (0.142) | (0.092) |
| | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] | [0.001] |
| Student observations | 1,077 | 471 | 606 | 313 | 764 | 272 | 805 |
| Control Mean | 3.525 | 3.562 | 3.497 | 3.267 | 3.631 | 3.418 | 3.560 |

Note: Note: The above estimates are derived from the follow-up survey conducted with 1,077 parents. The outcome variable in the above table is the average daily study hours in the past year as reported by parents in the follow-up survey. The specifications control for unbalanced covariate at the baseline—ownership of bicycle—and school-grade-gender fixed effects. Robust standard errors are clustered at the school-grade-gender level. BKY sharpened q-values to control for false discovery rate are indicated in square brackets. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels respectively.