Policy Research Working Paper 11286

Nudging at Scale

Evidence from a Government Text Messaging Campaign during School Shutdowns in Punjab, Pakistan

Koen Geven
Ayesha Tahir
Tazeen Fasih
Javaeria Qureshi
Amer Hasan
Sheena Fazili
Rabea Malik
Kevin MacDonald

WORLD BANK GROUP
Education Global Department
January 2026

Reproducible Research Repository

A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

Policy Research Working Paper 11286

Abstract

Text and voice messages have emerged as a low-cost and popular tool for nudging recipients to change behavior. This paper presents findings from a randomized controlled trial designed to evaluate the impact of an information campaign using text and voice messages implemented in Punjab, Pakistan during the COVID-19-induced school closures. This campaign sought to increase study time and provide academic support while schools were closed and to encourage reenrollment when they opened, to reduce the number of dropouts. The campaign targeted girls enrolled in grades 5 to 7. Messages were sent out by a government institution, and the campaign lasted from October 2020 until November 2021, when schools had permanently re-opened. Households were randomized across three treatment groups and a control group that did not receive any messages. The first treatment group received gender-specific messages that explicitly referenced daughters in their households, and the second treatment group received gender-neutral messages. A third group was cross randomized across the first two treatment arms and received academic support messages (practice math problems and solutions). The results show that the messages increased reenrollment by 6.0 percentage points approximately three months after the intervention finished. Gender neutral messages (+8.9 percentage points) showed larger effect size on enrolment than gender-specific messages (+ 4.3 percentage points), although the difference is not statistically significant. The message program also increased learning outcomes by 0.2 standard deviation for Urdu and 0.2 standard deviation for math. The paper finds a small positive effect on the intensive margin of remote learning and an (equivalent) small negative effect on the intensive margin of outside tutoring. In line with similar studies on pandemic remediation efforts, the paper finds no effect of the academic support intervention on learning. The findings suggest that increased school enrollment played a role in supporting the observed increase in learning outcomes.

This paper is a product of the Education Global Department. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at kgeven@worldbank.org. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.



The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Nudging at Scale: Evidence from a Government Text Messaging Campaign during School Shutdowns in Punjab, Pakistan*

Koen Geven[1], Ayesha Tahir[1], Tazeen Fasih[1], Javaeria Qureshi[2], Amer Hasan[1], Sheena Fazili[3], Rabea Malik[4], Kevin MacDonald[1]

Authorized for distribution by Fadila Caillaud, Practice Manager, Education Global Department, World Bank Group

Keywords: Education, Gender, Text Messages, Covid-19, Parenting
JEL classification: C93, I21, O15, J16

Cleared by: Fadila Cauillaud, Practice Manager, Education Global Practice, MENAAP

** The research team would like to express its sincere gratitude to the Government of Punjab, the Punjab School Education Department, the Project Management and Implementation Unit (PMIU) and the Punjab Information and Technology Board (PITB) and to research partners IDEAS, CERP, ITA and RCons for supporting the implementation of this research project. The team gratefully acknowledges generous funding from the Strategic Impact Evaluation Fund (SIEF) and the South Asia Region Gender Innovation Lab (SAR GIL) at the World Bank. The team also thanks participants at seminars organized by SIEF, SAR GIL, the Lahore University of Management Sciences (LUMS) and others for their feedback on the project. The team would like to thank Naveed Hussain for research assistance.*

[1] World Bank

[2] University of Illinois at Chicago, United States

[3] Independent researcher

[4] Lahore University of Management Sciences, Pakistan.

1. Introduction

In Pakistan, social norms regarding gender play a prominent role in households' schooling decisions (Jacoby & Mansuri, 2011). Policy makers see such norms as a major reason why only 57% of girls are enrolled in primary education, as compared to 64% of boys (Baron et al, 2024). Information can play an important role in conservative contexts like Pakistan because norm perception may not reflect actual behavior. Information can provide a quick pathway to updated beliefs about what constitutes normal behavior in such contexts, thereby changing the behavior of those receiving the information (Ali & Bénabou, 2020; Burstyn et al, 2020). Importantly, information can travel fast, at low cost and is thus scalable, in large part thanks to the wide spread of mobile phones (94% in the province under study, PSLM 2019/20).

The study leverages COVID-19-related school closures to assess the efficacy of a low-cost behavioral nudge in a context of strong social norms. COVID-19-related school closures led to concerns about potential learning losses very early in the pandemic period (Azevedo et al, 2020). Dropouts from school due to the closures were also prominent in several low-income countries (Sabarwal et al, 2023), with the overall level of dropouts observed estimated at 6 percentage points (Flor et al, 2022). Since then, empirical work has documented an average learning loss of around 0.14 standard deviation, with most studies documenting increases in educational inequality (Betthauser et al, 2023). School re-openings only partially recovered these learning losses, leading to continued concerns about how to build resilience within school systems (Dela Cruz et al, 2025). Girls were particularly at risk during the pandemic (Flor et al, 2022; Moscoviz & Evans, 2022), due to an array of factors including an observed increase in teen pregnancy (Zulaika et al, 2022), an increase in time allocated to household chores among girls and technology restrictions specific to girls (Kamran et al, 2022). Given low female labor force participation, parents in Pakistan also have rational incentives to underinvest in girls' education relative to boys' education (Aslam, 2009), leading them to potentially prioritize boys over girls when resources are constrained.

Well before the pandemic, a substantial empirical literature showed that simple, timely messages can be a cost-effective tool to shift education behaviors at low marginal cost. In high-income settings, personalized texts to students and families have raised college enrollment and persistence (Castleman & Page 2015, 2016), improved attendance and course completion (Kraft & Rogers 2015), and boosted achievement when messages translate broad goals into concrete, near-term actions (York, Loeb & Doss 2019; Doss et al. 2019; Bergman & Chan 2019). In middle-income contexts, text messages to parents have improved attendance, grades, and behavior (Berlinski et al, 2021; Hernández-Agramonte et al, 2024). Text messages have also been shown to have a direct effect on school attendance in Mozambique (De Walque & Valente, 2023). Reviews of the literature have found that these 'technology-enabled behavioral interventions' targeting informational barriers and behavioral inconsistencies can have a low to medium impact on learning, but with high cost-effectiveness (Rodriguez-Segura, 2023; Jordan, 2023). Related work from Pakistan shows that low-cost information interventions can reshape education choices and market outcomes (Andrabi, Das & Khwaja 2017).

Several studies have documented effective strategies to combat these learning losses, although sometimes with conflicting findings. While there is a growing literature on information interventions, some have found null effects on learning such as in Côte d'Ivoire (Wolf & Lichand 2023). There are also diverging findings on the effect of text messages on learning with some finding positive effects (Hernandez-Agramonte et al, 2024; Ome & Menendez, 2022), and others finding null effects (Angrist et al, 2022) or fade out effects (Schueler & Rodriguez-Segura, 2023). Even within pandemic-era

studies, impacts vary by design: messages that are generic, infrequent, or non-actionable can sometimes yield small or null effects, while more tailored, high-frequency outreach performs better, leaving open which framing works best for re-enrollment decisions (Angrist et al. 2022; Berlinski et al. 2021).

This paper makes three key contributions to the literature. First, the paper presents additional evidence on the impact of information campaigns in low-income settings. Second, given concerns about restrictions on girls at home, it explores the gendered dynamics of a text messaging campaign in a setting where social constraints on girls' education remain pertinent (Jacoby & Mansuri, 2011). Third, within the context of emerging literature on concerns about whether interventions can scale (Al-Ubaydli et al, 2019; Banerjee et al 2017), this paper estimates the effects of a program that is administered by a government, using government systems, rather than an independent implementing partner.

The intervention was an intensive information campaign combining text messages (SMS) and celebrity voice messages, implemented by a government institution aimed at parents of daughters who were enrolled in school prior to the pandemic. During the campaign, families received approximately seven text messages per week, and the campaign lasted from October 26, 2020, to November 20, 2021. Parents also received nine voice messages from Pakistani celebrities emphasizing the importance of schooling, re-enrollment and remote learning opportunities. These interventions were based on evidence that behaviorally-informed communication—when specific, regular, and actionable—can be more impactful than complex platforms, especially where connectivity is uneven and caregiver time is scarce (Escueta et al. 2020; Rodríguez-Segura 2022).

The intervention was tested in a randomized controlled trial conducted in select districts of Punjab, Pakistan, and targeted girls enrolled in grades 5-7 before the pandemic in a representative sample of public and private schools. Throughout the pandemic period, we followed a sample of 5,362 girls for which a baseline study was conducted in the summer of 2020. Families were interviewed at baseline and across three rounds of follow-up telephone surveys (rounds 1, 2 and 3), with the last round of data collection taking place between December 2021 and January 2022 after schools had been opened for approximately 3 months. We track enrollment and time use in the first two rounds, and we collect enrollment information and administer a learning assessment over the telephone in round 3, like similar research from the pandemic period (Crawfurd et al., 2021, Rodriguez-Segura and Schueler, 2022). We also conduct a qualitative strand of telephone interviews among a smaller sample of families in which we collect data on household dynamics during the pandemic period.

The intervention featured three treatment groups and a control group – with households randomly assigned across groups. The first treatment arm featured messages that used gender-specific language aimed at girls. The second treatment arm used gender-neutral language but was otherwise identical in content. For example, where the first treatment arm encouraged parents to support their *daughter*'s schooling, a typical message in the second treatment arm would frame the same around *children*'s or on *boys' and girls'* schooling. The idea behind these differences is that they would present this gender aspect as either salient or neutral. At the end of the summer of 2021 (when schools had reopened), families also received voice messages from national celebrities in Pakistan. In line with the intervention logic, Treatment 1 participants received a voice message that emphasized *girls'* schooling, while Treatment 2 participants received a voice message that emphasized *children's* schooling (in general). A third treatment arm (cross randomized across the first two groups) comprised of weekly text messages containing academic support using practice math problems (varying questions based on Subject

Learning Objectives of grades 5, 6, and 7) followed by their solutions to encourage and support children in learning Mathematics while at home during school closures. These messages were sent once schools had reopened in October 2021. The control group did not receive any messages from the program.

The paper finds strong evidence that the information campaign was effective (+6.0 p.p.) in increasing girls' enrollment in school once schools had permanently reopened. We observe that this effect is largest for students who are more likely to drop out, consistent with the idea that the effect is concentrated among students who are on the margin of deciding to drop out. The paper also finds strong evidence that an information campaign can increase test scores (+0.2 SD for Urdu, +0.2 SD for Math) after a long period of school closures. However, in line with Angrist et al (2022), we find that academic support messages (treatment 3) with practice math problems and their solutions do not have a statistically significant impact on learning outcomes.

The paper also explores potential pathways that may help us to understand how these effects came about. We find in rounds 1 and 2 of the surveys that text messages had a small positive effect on time spent on remote learning, while reducing the intensity of private tutoring by a similar amount, among those who participated in these activities. Our results suggest that re-enrolling in school is important to recover learning, which likely had been lost during the pandemic. Similarly, the results also suggest (but do not definitely prove) that gender-neutral messaging may be more effective in a context like Pakistan, as families do not seem to think that gender is a salient feature of their schooling decision. This finding however seems to align with the idea that gender-neutral policies may help girls as much or even more than gender-targeted interventions (Evans et al, 2022).

This paper proceeds as follows. In Section 2, we discuss the context of the intervention and provide details about the intervention itself. In Section 3, we describe the sample, attrition and the assignment to treatment and control groups, as well as the qualitative work to better understand the lived experience of families during school closures. In Section 4, we provide the results of treatments on re-enrollment and learning levels, the heterogeneity analysis, checks on the analytical sample and the results of the treatments on time-use. Section 5 discusses the results of the qualitative strand of work. Finally, in Section 6, we discuss the results and provide suggestions for further research and policy.

2. Context and Intervention

2.1. Context

Punjab's school system is vast: in 2019–20 the province hosted more than 52,000 public schools and 45,000 private schools from primary to higher secondary education, together enrolling over 22 million children (Pakistan Institute of Education, 2023). The system also grapples with the significant challenge of over 6 million out-of-school children, which constitutes 26% of the province's population aged 5-16 (PSLM 2019/20). Altogether, this makes Punjab's basic and secondary education system one of the largest in the world.[5] The education system in the province has attracted sustained attention in the economics literature (Andrabi et al, 2008; Andrabi et al, 2023). These studies underscore that

[5] To put that in perspective, Punjab alone educates more children than entire countries such as England (≈9 million students) or large states such as California (≈6 million public K–12 students).

Punjab's education sector is not only unusually large but also unusually privatized, making it an important case for global debates on school systems and education policy.

Pakistan had a very long period of school closures during the COVID-19 pandemic period – and official estimates suggest that school closures lasted for 37 weeks of full closures and 24 weeks of partial closures, leading to 245 learning days lost in total (UIS 2022) . All schools in Punjab were closed by government order on March 15, 2020, to prevent further spread of the COVID-19 virus. Between April 1, 2020, and April 9, 2020, the country was in lockdown to stop the spread of the pandemic. Due to the cyclical nature of the COVID-19 spread, school closures followed an intermittent pattern in which schools were often opened temporarily or intermittently (with children attending on alternate days for physical distancing). Schools were temporarily reopened between September 15, 2020, and November 26, 2020. After the second lockdown, schools were reopened from January 18, 2021 (High Schools) or February 1, 2021 (Primary Schools) until March 15, 2021. Schools then had summer break until August 1. A short final period of school closures was imposed from September 11, 2021. Finally, schools were permanently opened on September 16, 2021.

Table 1: School closures and re-openings during the intervention period

School Closure Date	School Reopening Date
March 15, 2020	September 15, 2020
November 26, 2020	January 18, 2021 (*some districts: February 1, 2021*)
March 15, 2021 (*some districts March 26, 2021*)	August 1, 2021
September 11, 2021	September 16, 2021

2.2 Intervention Design

On request of the Government of Punjab, the research team (in collaboration with the provincial government and local stakeholders) designed the information campaign to support girls' learning during the period of school closures and to nudge families to re-enroll their daughters after schools re-opened. Household statistics revealed that at the start of the pandemic, 94% of households in Punjab province had access to a mobile telephone (PSLM 2019/20), making this a feasible strategy to reach households.

The main intervention included a set of text messages (SMS) that were sent to parents in a language that would be easy to understand for most parents. Table 2 provides an overview of the main treatment arms of the intervention. The main intervention was divided into two treatment arms. The first treatment arm featured messages that were gender-specific in tone. The second treatment arm used gender-neutral language but was otherwise identical in content.

The content of these messages included behavioral nudges around engaging with children's learning in general and during remote learning; and to encourage reenrollment once schools re-open. The SMS contained a message with a maximum of 160 characters that would fit on a typical mobile phone screen (regardless of whether that was a smartphone or not). Text messages were written in English by the research team, and then translated into Urdu, the national language of Pakistan and the official language of instruction. All translations were reviewed for consistency and clarity through multiple rounds of review, including by local stakeholders.

The content of the messages was developed based on recommendations from the literature on text message nudges (York & Loeb, 2018; Minero, 2014; Cabell et al, 2019). Messages included a "Fact" text that was sent at the start of the week and informed families of the skill of the week and its importance for the academic growth of their child. For example, "Fact: Research shows that reading daily can significantly improve children's academic performance - We learn from practicing daily." This message was followed by two to three 'Action-' or 'Talking Point-' based messages. For example, "Action: Find newspapers, magazines, or brochures to help your daughter practice Urdu reading skills. Ask her to read a passage from the newspaper/magazine to you."

Voice messages were recorded by Pakistani celebrities[6] and were sent out by the government across Treatment 1 and 2 arms between August 24 and September 2, 2021, as schools in Punjab had reopened. These messages were informed by similar work in India on pandemic health information campaigns (Banerjee et al, 2020). The content of voice messages covered themes such as the importance of schooling, re-enrollment and remote learning opportunities. The content of voice messages emphasized girls and daughters for Treatment 1 participants, and featured gender-neutral language for Treatment 2 participants. A sample voice message was 'This is Shehzad Roy (*or other celebrity name*) and I bring to you an important message from the Government of Punjab. School closures due to COVID-19 have disrupted our girls (*or children's*) education and threatened their learning. Our girls (*or children*) are facing difficulties keeping up with learning goals. You as parents can play a crucial role in securing your girls' (*or children's)* futures. I request you to make sure your girls (*or children)* return to school, now that schools are open. You can ensure a bright future for your girls (*or children)*, your family and the country. Take a new step toward a better tomorrow for your family and Pakistan.'

Given the risk of learning loss due to the pandemic-related school closures (Azevedo et al, 2021), we also designed a third treatment arm comprised of weekly text messages containing several simple math problems (varying questions based on Subject Learning Objectives (SLOs) of grades 5, 6, and 7) followed by their solutions to encourage and support children in learning Mathematics while at home during school closures. This intervention was designed in collaboration with Idara-e-Taleem-o-Agahi (ITA), a nonprofit institution in Pakistan which had piloted a similar activity for out-of-school children in Punjab. Mathematics problems were tailored for children in different grades in the sample and curriculum experts designed the learning exercises. Messages covered content that would typically feature in the grade in which they were enrolled at the time of the intervention.

Table 2: An overview of the treatment arms in the intervention

	Content	**Example message**	**Total number of messages sent *(number of weeks)***
Treatment 1 (T1)	**Girl-targeted** text messages focused explicitly on girls' enrollment, learning, and engagement with	FACT: Your daughter could fall behind if they do not revise their daily school lessons	119 text messages + 9 voice messages *(16 weeks)*

[6] Celebrity voice messages were recorded by Shehzad Roy (Singer), Moin Khan (Cricketer), Shoaib Malik (Cricketer), Sayein Zahoor (Singer), Hira Mani (Actress), Humera Arshad (Singer), Javed Sheikh (Actor), Mustafa Qureshi (Actor), Sarfraz Ahmed (Cricketer), Shazia Manzoor (Singer), Sheraz Uppal (Singer), Anwar Ali (Cricketer), Rumman Raees (Cricketer), Rambo (Actor), Arif Lohar (Singer), Sahiba (Actress), Nisho Begum (Actress), Honey Albela (Comedian) and Murad Raas (Minister of Education, Punjab province, Pakistan, at the time of the intervention).

	distance learning programs.	حقیقت: اگر آپ کی بیٹی سکول میں پڑھے ہوئے اسباق کی مشق نہیں کرے گی تو وہ باقی طلبہ سے پیچھے رہ جائے گی۔	
Treatment 2 (T2)	**Gender-neutral** messages focused on all children in the household.	FACT: Your child could fall behind if they do not revise their daily school lessons حقیقت: اگر آپ کا بچہ سکول میں پڑھے ہوئے اسباق کی مشق نہیں کرے گا تو وہ باقی طلبہ سے پیچھے رہ جائے گا۔	119 text messages + 9 voice messages *(16 weeks)*
Treatment 3 (T3)	**Academic Support** sharing math exercises and accompanying solutions with parents for children (in grades 5-7) as per Subject Learning Objectives (SLOs) in each grade through text messages.	Question: If the cost of 15 balls is Rs. 300, then find the cost of one ball? Answer: This is a division question. Dividing 300 by 15 gives the cost of one ball as Rs. 20. سوال :اگر15گیندوں کی قیمت 300روپے ہو تو ایک گیند کی قیمت معلوم کریں؟ جواب :یہ تقسیم کا سوال ہے۔ 300کو15پر تقسیم کرنے سے ایک گیند کی قیمت حاصل ہوگی 20روپے۔	Grade 5: 84 text messages Grade 6: 94 text messages Grade 7: 89 text messages *(4 weeks)*
Control Group	Did not receive any messages		

The target population for the intervention was households with a girl enrolled in grades 5 to 7. Text messages were sent to parents' phone numbers that were supplied by the research team from records obtained from a school survey and government records (see sample section). All nudges were sent from an official government account, which was easily identifiable for parents. All text messages were reviewed by the School Education Department before finalization and then sent for delivery to the Punjab Information Technology Board, a government department focused on Information Technology, which had coordinated similar official campaigns before. Phone numbers were tested and deemed functional by the relevant government institutions before the initiation of the messaging campaign.

In total, households in treatment received (Table 2) a minimum of 119 and a maximum of 213 messages (depending on the treatment arm), which breaks down as 119 in Treatment Arms 1 and 2 each, and 84-94 messages in Treatment Arm 3 (depending on the grade the child was enrolled). For Treatment Arms 1 and 2, this represented 7.4 messages in a week when messages were received. For Treatment 3, this represented 21-23.5 messages per week. Messages in Treatment Arm 1 and Treatment Arm 2 were sent during 16 weeks between October 16, 2020 and September 9, 2021; text messages for Treatment Arm 3 were sent in a four-week period between October 25, 2021 and November 20, 2021.

While information campaigns are low-cost, this intervention was on the higher side of costs for such interventions. Cost data was collected through all stages of the intervention from the government institutions, and all partners involved in its implementation (including the World Bank). The marginal cost of a text message in Pakistan is extremely low, at about a quarter of a cent in US dollar terms. Given the large volume of text and voice messages and overheads involved in the coordination and participation of the government and partners, the total marginal cost of the intervention was US$ 13.84 per treated child. When aggregating overheads, however, the total cost (to all stakeholders) of the intervention was US$ 48 (Treatment 1), US$ 49 (Treatment 2) and US$ 30 (Treatment 3) per treated child.

3. Sampling, Randomization, Data and Empirical Strategy

3.1 Sample

The sample for this survey was drawn from two sources. The first source was a school survey that was representative of public and private schools in Punjab. This school survey had been conducted in early 2018 and included telephone numbers of households for a subset of students. At the start of the present study, in 2020, these students were enrolled in grades 5 to 7. A random sample of 5,000 girls was intended to be selected from this group for administration of the baseline survey. Due to a high rate of non-response from these households, the resulting sample size was 2,774 girls. To increase sample size, the Punjab government supplied an additional list of 93,990 students and their household's phone numbers from 518 public schools in Muzaffargarh and Rahim Yar Khan districts – two of the six districts in the 2018 school survey, and two of the poorest districts in the province. A sample of four girls was randomly selected from each school in this data source. This resulted in an additional sample of 2,588 girls, for a total of 5,362 girls for which a baseline study was conducted. We discuss the data that we collected on this sample, attrition from the survey, and the statistical model that we apply to the data in the next sections.

3.2 Data

Data for assessment of the interventions was collected through computer assisted telephone interviews (CATI) and included a baseline survey, and three rounds of follow-up surveys. In addition, three rounds of qualitative telephone surveys were also conducted to understand the lived experiences of parents and daughters during the COVID-19 pandemic period and strengthen the understanding of the empirical findings.

The baseline survey was conducted during August and October 2020. The baseline survey collected data on child age, household size and household characteristics including assets owned, as well as information about time-use during the pandemic period.

Rounds 1 (February-March 2021) and 2 (September-October 2021) were conducted during periods when schools had temporarily reopened and captured whether the students had been reenrolled in school. The questionnaires also included time use across common activities affecting learning. The final round (Round 3, December 2021-January 2022) took place once schools had permanently re-opened. The Round 3 survey measured learning outcomes as measured by a short phone-based assessment for Urdu (the national language of Pakistan) and Mathematics. Table 3 shows the six outcomes of interest that are available and reported in the paper (Table 3).

Table 3. Outcome variables defined by survey round

	Included in quantitative survey round		
	1	**2**	**3**
Re-enrollment in school	X	X	X
Learning Outcomes			
Phone-based Urdu assessment score			X
Phone-based Math assessment score			X
Time use			
Remote learning	X	X	
Other academic learning	X	X	
Household chores	X	X	

Note: For time use variables, three outcomes were defined: whether any time was spent on the activity, the number of minutes spent, and the number of non-zero minutes spent.

In parallel to the CATI-based household surveys, the qualitative stream of work was designed to provide an understanding of day-to-day experiences during school closures through the voices of girls, mothers and fathers. It also served to provide understanding of how the participants were interacting with the interventions introduced to address learning loss due to school closures. Qualitative interviews were conducted via telephone by trained researchers in three different instances.[7] The first set of qualitative interviews was conducted following the baseline, and comprised 30-minute telephone interviews with mothers and daughters, A second set of interviews was conducted following the midline and comprised 30-minute telephone interviews with mothers and daughters and a subset of fathers, and a third set of interviews was carried out following the endline, and also comprised 30-minute telephone interviews with mothers and daughters. The interviews asked both about steps taken by government institutions to allow the continuity of education during school closures (such as education channels on TV) as well as the messages that were sent under this evaluation. A purposive sampling strategy was adopted to ensure representation across treatment arms, districts, socio-economic strata, and past academic performance (measured in the 2018 survey). For this stream of data collection, 72 households participated across 6 districts, 36 from treatment arms and 36 from control.

3.3 Randomization

The target population to measure the impact of the treatments were girls enrolled in grades 5 to 7 in Punjab province in 2020. After the baseline interviews, sampled girls were randomly assigned to one of three groups: a control group (C), Treatment 1 (T1), and Treatment 2 (T2). Treatment 3 (T3) was cross-randomized across the Treatment 1 and 2 groups. Assignment was randomized within each school. Because the girls in original school survey had test score data from 2018, this was used as an implicit stratification variable in which girls were sorted by test score, the girl with the highest test

[7] Researchers were based at and trained by IDEAS Pakistan.

score was randomly assigned to one of the three groups (T1, T2 and C) and subsequent girls were assigned systematically. For the sample from the list of children supplied by the government, girls were randomly assigned to one of the three arms, but no additional stratification was applied.

3.4 Analytical Sample, Balance and Attrition

Due to the challenges of administering surveys during the pandemic, attrition from the sample was substantial. Table 4 shows the evolution of the sample by survey round. By the last round of the survey, the sample included 2,475 girls on which data was collected. Out of this sample, 2,098 respondents completed the full Round 3 survey, and these comprised the analytical sample. There was also complete data in the Baseline and the Round 1 survey for this same sample, although there was a slight dip in participation of the analytical sample in round 2.

Table 4. Sample Attrition by Survey Round

	Baseline	**Follow-Up**		
	(August – October 2020)	**Round 1** (February – March, 2021)	**Round 2** (September – October, 2021)	**Round 3** (December 2021 – January 2022)
Total Sample				
Total N	5,362	4,080	3,536	2,475
(row percentage)	*(100%)*	*(76%)*	*(66%)*	*(46%)*
Any Treatment	3,571	2,733	2,349	1,623
(row percentage)	*(100%)*	*(77%)*	*(66%)*	*(45%)*
Treatment 1	1,789	1,380	1,179	824
(row percentage)	*(100%)*	*(77%)*	*(66%)*	*(46%)*
Treatment 2	1,782	1,353	1,170	799
(row percentage)	*(100%)*	*(76%)*	*(66%)*	*(45%)*
Treatment 3 (cross randomized)	797	684	636	797
(row percentage)	*(100%)*	*(86%)*	*(80%)*	*(100%)*
Control	1,791	1,347	1,187	852
(row percentage)	*(100%)*	*(75%)*	*(66%)*	*(48%)*
Analytical Sample				
Total N	2,098	2,098	1,712	2,098
(row percentage)	*(100%)*	*(100%)*	*(82%)*	*(100%)*
Any Treatment	1,377	1,377	1,127	1,377
(row percentage)	*(100%)*	*(100%)*	*(82%)*	*(100%)*
Treatment 1	708	708	585	708
(row percentage)	*(100%)*	*(100%)*	*(83%)*	*(100%)*
Treatment 2	669	669	542	669
(row percentage)	*(100%)*	*(100%)*	*(81%)*	*(100%)*
Treatment 3 (cross randomized)	680	680	567	680
(row percentage)	*(100%)*	*(100%)*	*(83%)*	*(100%)*
Control	721	721	585	721
(row percentage)	*(100%)*	*(100%)*	*(81%)*	*(100%)*

Notes: Row percentages displayed to facilitate comparison of attrition across subsamples. Any treatment is the sum of Treatment 1 and 2. Treatment 3 is cross-randomized across treatment 1 and 2. The analytical sample is the sample that we observe in round 3 (where we collect the main outcome variables) and for which we observe two background variables that are unbalanced in the sample (see Table A2-A4) and the discussion below attrition in this section of the paper.

Simple t-tests confirm that there was no differential attrition in the analytical sample by treatment group (Table 5). In the group receiving Any Treatment, 61.4 percent dropped out from our Analytical Sample, 60.4 percent for Treatment 1, and 62.5 percent in the group receiving Treatment 2. In the control group, 59.7 dropped out from the Analytical Sample. In the group receiving Treatment 3 (which was cross-randomized across Groups 1 and 2 in the summer of 2021), 14.7 percent dropped out from the main sample, versus 15.4 percent in the Control group. Overall, while there is slightly higher attrition in the treatment groups (as compared to the control groups), this difference is not statistically significant except in the difference between Treatment 2 and the Control group with a p-value of 0.096. But this difference disappears in regression analysis, controlling for baseline observables (see Appendix Table A.1).

Table 5: Sample Attrition between Baseline and Analytical Sample

	Mean analytical sample among treatment	Mean analytical sample among control	Difference	p-value of the difference	Total Observations
Any Treatment	0.614	0.597	0.017	0.230	5,362
Treatment 1	0.604	0.597	0.007	0.677	3,580
Treatment 2	0.625	0.597	0.027	0.096	3,573
Treatment 3	0.147	0.154	0.007	0.693	1,649

Note: difference in means between treatment and control group estimated through simple t-test. Regression analysis in Appendix Table A.1.

Balance in background characteristics including attrition rates was estimated for all combinations of treatment and control groups and are presented in Appendix Tables A2-A4. Balance was estimated by estimating equation (1) with the sample restricted to the treatment group in question and the control group.

$$Y_i = \beta_0 + \beta_1 T_i + e_i \quad (1)$$

where Yi denotes the background variable of interest and T_i the treatment (varying from Treatment 1-3). Fixed effects at the school level were included in the model. Estimates were weighted by observations' inverse selection probability and standard errors were adjusted to account for the sampling design including clustering at the school level. The only variables that are unbalanced in the final analytical sample are the total number of children in the household and mother's education in the comparison of Any Treatment versus Control).[8] We control for these variables in the preferred specification of the analysis.

An overview of the analytical sample, with means and standard deviations of key variables, is included in Table 6.

[8] There is also imbalance in the number of children in the household in the comparison of Treatment 1 versus Control. In other balance tables, there is no statistically significant difference in any of the observed variables.

Table 6. Analytical Sample Descriptives

	Mean value (S.D.)	N
Female	1.000 (0.000)	2,098
Age <13 (Tercile – 1)	0.492 (0.500)	2,098
Age 13 (Tercile – 2)	0.240 (0.427)	2,098
Age > 13 (Tercile – 3)	0.268 (0.443)	2,098
Grade enrolled at baseline	5.936 (0.770)	2,024
Rural School	0.743 (0.437)	2,098
Asset Tercile - 1	0.352 (0.478)	2,098
Asset Tercile - 2	0.314 (0.464)	2,098
Asset Tercile - 3	0.331 (0.471)	2,098
Number of Adults in the Household	3.831 (2.401)	2,098
Number of Children in the Household	3.010 (1.318)	2,098
Father Not Educated	0.287 (0.452)	2,098
Father's Education - Less than Grade 10	0.427 (0.495)	2,098
Father's Education - Grade 10 and above	0.286 (0.452)	2,098
Mother Not Educated	0.510 (0.500)	2,098
Mother's Education - Less than Grade 10	0.352 (0.478)	2,098
Mother's Education - Grade 10 and above	0.137 (0.344)	2,098
Any Treatment	0.656 (0.475)	2,098
Treatment 1 (Girl-Specific)	0.495 (0.500)	1,429
Treatment 2 (Gender-Neutral)	0.481 (0.500)	1,390
Treatment 3 (Academic Support)	0.485 (0.500)	1,401
Control Group	0.344 (0.475)	2,098
Participated in Round 1 Survey	1.000 (0.000)	2,098
Participated in Round 2 Survey	0.816 (0.388)	2,098
Participated in Round 3 Survey	1.000 (0.000)	2,098

Notes: descriptive statistics displayed for analytical sample only. Sample for Treatments 1, 2 and 3 differs, given exclusion of children in other Treatment groups (Treatment 1 vs. Control = 1,429 children, Treatment 2 vs. Control = 1,390 children, Treatment 3 vs. Control 1,401 children).

3.5 Statistical Model

To estimate the impact of assignment to treatment, we estimate the following equation:

$$Y_{ist} = \beta_0 + \beta_1 T_{is} + X_{is} + \mu_{is} + X_{is} + \epsilon_{ist} \quad (2)$$

Each outcome Y_{ist} for child i in school s measured at time t is regressed on a dummy for treatment assignment, T_{is}, and randomization strata fixed effects, μ_{is}. We also include a vector of covariates, X_{is}, that are a baseline measure of the outcome and additional controls for baseline variables where imbalance was detected among treatment and control groups at baseline.

$$Y_i = \beta_0 + \beta_1 T_{1+2i} + X_{is} + \alpha_i + e_i \quad (3)$$

Each equation was estimated with a restricted sample that included only observations in the treatment group of interest and the control group. The second equation included the treatment of interest and school fixed-effects and a set of control variables for which imbalance was found or suspected (see discussion above).

$$Y_i = \beta_0 + \beta_1 T_{1i} + \beta_1 T_{2i} + X_{is} + \alpha_i + e_i \quad (4)$$

A third pair of equations were also estimated which included both treatments in the same model. This enabled tests for whether there were any differences in effect size between treatments. These equations were estimated using OLS with sample weights based on the inverse selection probability of the observations and standard errors were adjusted for clustering at the school level.

4. Results

Table 7 shows the effects of the program at endline on enrollment, and Table 8 shows the effects of the program on Urdu and Math test scores (Panel A and B respectively). Column 1 reports the effect of the information campaign relative to the control group, whereas Columns 2 and 3 show the effect of the message treatment based on whether the students received gender-neutral messages or gender-specific messages separately. Column 4 shows the effect of the Academic Support intervention.[9] In table 8, column 5 shows effects on test scores controlling for girls being enrolled in school during R3 data collection. The results presented here are intent-to-treat effects on whether households were randomly assigned to treatment.

Table 7: Whether Child Re-Enrolled at Round 3 (Endline)

	(1)	(2)	(3)	(4)
Any Treatment	0.060**			
	(0.030)			
Treatment 1 (Gender-Specific)		0.043		
		(0.033)		
Treatment 2 (Gender-Neutral)			0.089**	
			(0.038)	
Treatment 3 (Academic Support)				0.069*
				(0.038)
Mean value of re-enrollment in control group	*0.926*	*0.926*	*0.926*	*0.926*
R^2	0.54	0.66	0.62	0.60
N	2,098	1,429	1,390	1,401

Notes: Each row reports the results of a different model. All models control for imbalanced variables at baseline (household size and mother's education) and use school fixed effects. Results are weighted. Standard Errors are denoted in parentheses. * p<0.1; ** p<0.05; *** p<0.01

The results in Table 7 show that the information campaign significantly increased girls' enrollment at endline. Girls in households who received the messages were 6 percentage points more likely to be enrolled in school compared to girls in the control group upon reopening of schools (above 92.6

[9] Note that while the academic support message intervention was implemented after schools had already re-opened, it is still plausible that this intervention would have causally effected enrollments. Enrollment in Pakistan is a continuous process, in which students often re-enroll at a later stage in the academic year.

percent in the control group). In other words, the message program effectively led to re-enrollment in almost the entire treatment group. The results in Columns 2 and 3 suggest that the gender-neutral message program significantly increased enrollment by 9 percentage points and the gender-specific message program increased enrollment by 4 percentage points, though the latter effect is not statistically significant. While the effect sizes displayed here suggest that the difference is driven by the gender-neutral messages, the difference between the estimates of Treatment 1 and Treatment 2 is not statistically significant (the difference is -0.046, S.E. 0.036, p-value 0.204). We also find that the academic support (sample math exercises) increased re-enrollment by 6.9 percentage points and this effect is statistically significant.

The results in Column 1 of Table 8 (Test Scores by Round 3) suggest that information campaign significantly increased overall test scores by 0.203 and 0.201 standard deviations in Urdu (Panel A) and Mathematics (Panel B), respectively. In line with the pattern of enrollment effects, we find that the effect of the gender-neutral program on both Urdu and Math scores seems larger than the effect of the gender-specific messages which is also not statistically significant. However, these differences between Treatment 1 and Treatment 2 are not statistically significant for either Urdu (the difference is -0.128, S.E. 0.135, p-value 0.344) or Math (the difference is -0.177, S.E. 0.150, p-value 0.238). The effect of Treatment 3 (Academic Support messages) on test scores is not statistically significant.

In the last column, we explore a potential pathway through which the effects on test scores are related to re-enrolling. We find that the test score for Urdu shows an increase of 0.179 standard deviation even after controlling for re-enrollment in R3 and remains statistically significant, whereas the test score for math remains positive but loses statistical significance. We also see that re-enrollment by Round 3 has a substantial independent effect on test scores for both Urdu (0.4 standard deviation, statistically significant) and Math (0.503 standard deviation, statistically significant). Taking the point estimates at face value suggests that about 12%-15% of the positive treatment effect on Math and Urdu scores is attributable to the fact that the program increases enrollment rates. While this is not a direct test of this pathway, this may imply that part of the observed effect on test scores is produced by re-enrolling in school.

Table 8: Test Scores at Round 3 (Endline)
A. Urdu

	(1)	(2)	(3)	(4)	(5)
Any Treatment	0.203**				0.179**
	(0.092)				(0.091)
Treatment 1 (Gender-Specific)		0.169			
		(0.111)			
Treatment 2 (Gender-Neutral)			0.296***		
			(0.109)		
Treatment 3 (Academic Support)				0.154	
				(0.116)	
Child was re-enrolled at Round 3					0.400**
					(0.181)
R^2	0.64	0.72	0.69	0.71	0.65
N	2,098	1,429	1,390	1,401	2,098

B. Mathematics

	(1)	(2)	(3)	(4)	(5)
Any Treatment	0.201*				0.171
	(0.107)				(0.105)
Treatment 1 (Gender-Specific)		0.130			
		(0.135)			
Treatment 2 (Gender-Neutral)			0.308**		
			(0.121)		
Treatment 3 (Academic Support)				0.130	
				(0.126)	
Child was re-enrolled at Round 3					0.503**
					(0.210)
R^2	0.63	0.70	0.71	0.71	0.64
N	2,098	1,429	1,390	1,401	2,098

Notes: All models control for imbalanced variables at baseline (household size and mother's education) and use school fixed effects. Results are weighted. Standard Errors are denoted in parentheses. * p<0.1; ** p<0.05; *** p<0.01. Means of dependent variable in the control group omitted, because test scores are standardized with mean zero.

Next, we explore heterogeneity in these effects by predicted dropout risk following the approach in Edmonds et al. 2021. Specifically, we predict control group dropout at endline using baseline characteristics and we use this model to assign predicted dropout propensities to all girls in our sample based on their baseline characteristics. In Table 8 we show heterogeneity in the estimated effects of either treatment on outcomes by predicted dropout risk. We characterize girls as having low dropout risk (measured as those with below median predicted dropout propensity) and high dropout risk (measured as those with above median predicted dropout propensity) and report the main effect of either treatment as well as the interaction effect of either treatment with the indicator for high dropout risk.

The coefficient on the indicator for above median dropout propensity in Table 9 shows that being in the high dropout risk group is associated with significantly lower likelihood of enrollment at endline. It is also associated with lower Urdu and Math test scores though these coefficients are not statistically significant. Turning to the effects of our program, we find that the information campaign has no significant effect on enrollment for girls with low dropout risk but that the effect is significantly larger for girls with high dropout risk.[10]

The heterogeneity of the test score effects is consistent with the pattern of heterogeneity we find for enrollment effects, i.e., the point estimates suggest larger benefits of the treatment for girls with high dropout risk. However, the estimated coefficients of the main treatment effect as well as the interaction terms lack statistical significance when studying test scores. Overall, our efforts to understand heterogeneity in our results are hampered by imprecision, but the results for enrollment and the point estimates for the other outcomes suggest that our treatment effects in the full sample might be driven by relatively disadvantaged students who are at higher dropout risk ex ante.

Table 9: Heterogeneity analysis for re-enrollment at Round 3

[10] We separately confirm that the effect of the program on enrollment for the high-risk group is statistically different from zero.

	Re-Enrollment	Urdu	Math
Above Median Propensity Score x Either Treatment Received	0.209***	0.261	0.234
	(0.073)	(0.225)	(0.239)
Any Treatment	-0.030	0.091	0.090
	(0.020)	(0.149)	(0.161)
Above Median Propensity Score	-0.201***	-0.257	-0.144
	(0.058)	(0.175)	(0.192)
Mean value of re-enrollment in control group	*0.926*		
R^2	0.56	0.65	0.63
N	2,098	2,098	2,098

Notes: We predict children's propensity to dropout at endline for the sample of girls in the control group by using a range of variables such as terciles of age, number of children in the household, asset terciles, parents education, district, school type, household location (urban/rural), whether the household faced any shocks such as loss of income during the pandemic; dummies for missing data in any of the variables; and full interactions of these. We have a total of 721 children in the control group at round 3. Age terciles 1-3 include 358, 164 and 199 children respectively. Asset terciles 1-3 include 272, 227 and 219 children and a dummy variable for missing asset score includes 3 children. In contrast, those who received any treatment include 1,377 children. We then use an indicator for above median propensity score along with a dummy for treatment and an interaction of the two to see how dropout propensity affects outcomes – enrollment at round 5 and assessment scores. All models control for imbalanced variables at baseline (household size and mother's education) and use school fixed effects. Results are weighted. Standard Errors are denoted in parentheses. * p<0.1; ** p<0.05; *** p<0.01***

As a consistency check on our results, Table 10 shows the effect of the information campaign for the sample of students we can consistently follow through all the rounds on test scores (Panel A) and re-enrollment (Panel B). Note that this sample is somewhat smaller than our analytical sample, because we had higher attrition in round 2 of the survey (see Table 4). As in our main results in Table 7, the information campaign had a positive effect on Urdu and Math test scores, but the effects are slightly smaller than in our main results and lack statistical significance. While we only tested the students once at endline, we have information on enrollment in each of the survey rounds so we also assess how the effect on enrollment evolved over time in Panel B (again for the sample that consistently participated across rounds). Columns 1 and 2 show that the effect of the information campaign in rounds 1 and 2 is positive but smaller and statistically insignificant. At endline, we find that any treatment results in a 7.6 percentage point increase in enrollment for the sample of girls we are able to consistently follow across all survey rounds.

Table 10: Treatment effects among sample that consistently participated in all survey rounds

A. Test Scores by Round 3			
	Urdu Scores	Math Scores	
Any Treatment	0.190	0.119	
	(0.120)	(0.124)	
R^2	0.67	0.67	
N	1,712	1,712	
B. Re-enrollment across survey rounds			
	Round 1	Round 2	Round 3

Any Treatment	0.049	0.057	0.076**
	(0.033)	(0.037)	(0.037)
Mean value of re-enrollment in control group	*0.924*	*0.920*	*0.939*
R^2	0.64	0.66	0.63
N	1,712	1,712	1,712

Notes: All models control for imbalanced variables at baseline (household size and mother's education) and use school fixed effects. Results are weighted. Standard Errors are denoted in parentheses. * p<0.1; ** p<0.05; *** p<0.01. Means of dependent variable in the control group are omitted for test scores, because these are standardized with mean zero.

To explore the mechanisms behind the observed effects, we estimate the effects of the intervention on time-use and present the results on both the extensive (whether the activity was undertaken) and intensive margins (the amount of minutes, and the amount of non-zero minutes spent on the activity in the 7 days before the survey) in Table 11. Note that we only observe time-use in Rounds 1 and 2 of the survey, but not in Round 3. We observe a small effect on the intensive margin in the case of weekly remote learning (an increase of 7.5 minutes per week) in the first round of the survey, which is offset by a reduction in time spent with an outside tutor (a decrease of 11.8 minutes per week).

There is no difference in the extensive margin of either remote learning or tutoring, however. This may be explained by the emphasis of the early text messages on engaging in the remote learning programs that were offered by the government, as well as a promotion of health behaviors including social distancing (i.e. causing a reduction in time spent with a tutor, which is typically offered face-to-face). By the second round, however, these effects have already faded out. The treatments also did not have any effect on either the extensive or the intensive margin of other types of learning (e.g. school attendance when schools were open, reading), or engagement in household chores in either survey round.

Table 11: Treatment effects on time-use in the 7 days before survey administration, Round 1 and Round 2

	Minutes, R1	Non-Zero Minutes, R1	Any, R1	Minutes, R2	Non-Zero Minutes, R2	Any, R2
Remote Learning						
Any Treatment	1.097	7.514**	0.012	0.771	-13.098	0.017
	(1.041)	(2.895)	(0.026)	(1.978)	(10.980)	(0.033)
Mean in the control group	2.578	19.836	0.130	3.671	52.213	0.070
R^2	0.72	0.99	0.78	0.61	0.97	0.61
N	1,566	108	1,566	1,631	168	1,631
Other learning						
Any Treatment	2.043	1.530	-0.008	9.453	-1.430	0.069
	(7.099)	(6.529)	(0.036)	(10.943)	(5.827)	(0.061)

Mean in the control group	111.315	123.602	0.901	109.354	131.449	0.832
R^2	0.65	0.72	0.52	0.62	0.65	0.67
N	1,566	1,415	1,566	1,631	1,425	1,631
Household Chores						
Any Treatment	-1.489	-7.630	0.042	0.443	4.491	-0.002
	(4.426)	(6.455)	(0.074)	(3.471)	(5.985)	(0.055)
Mean in the control group	30.516	52.789	0.578	30.931	55.349	0.559
R^2	0.70	0.80	0.71	0.69	0.81	0.67
N	1,566	882	1,566	1,631	885	1,631
Tutoring						
Any Treatment	-1.657	-11.801**	0.030	-5.945	-9.084	-0.005
	(9.240)	(5.869)	(0.068)	(10.047)	(7.291)	(0.062)
Mean in the control group	41.930	133.054	0.315	35.619	136.335	0.261
R^2	0.70	0.87	0.67	0.69	0.86	0.69
N	1,566	442	1,566	1,631	392	1,631

(1) Any Treatment Effect on time-use All models control for imbalanced variables at baseline (household size and mother's education) and use school fixed effects, Results are weighted. Standard Errors are denoted in parenthesis* p<0.1; ** p<0.05; *** p<0.01***

5. Findings from Qualitative Research

During implementation of the study, we conducted a series of qualitative interview rounds to better understand how messages were perceived and shared within households, and to have a richer understanding of the constraints on households and children's participation in schooling and remote learning opportunities. These interviews provide clues to better understanding of the pathways that we observe above.

Interviewees noted that fathers were usually in control of the telephone in the household, which could provide a bottleneck between the text message and the rest of the family. For several respondents, access to the phone, parent's literacy level and household economic conditions impacted the extent to which families interacted with the text messages. Nevertheless, fathers often did ensure that messages would be shared within the household. Interviewed fathers acknowledged receipt of the messages and how they would let their wife and children read the messages in the evening when they would return from work. As one father recounts: *"They [the messages] were mostly regarding children and their education. I don't really remember the details of these messages. I went home and made my wife and daughter read them. I made sure that she read each and every message. Maybe she remembers but I don't as she's better at understanding both English and Urdu. The messages stated that nothing should be more important than our children's education. We should focus on their education, make them watch educational TV channels and make sure they are studying every day.*

I never found them to be problematic or considered them a nuisance. Even when I got 10 messages in a day I made sure my daughter read all of them." None of the interviewed fathers were bothered by the text messages, showing that official government communication can be regarded as a positive channel of support.

The qualitative interviews confirmed that nation-wide and local lockdowns led families to be housebound for up to three months, leading to both income losses and shifting the gendered burden of work. All mothers and daughters interviewed for this qualitative study reported an increase in daily household chores during lockdown, including increased cleaning, cooking, washing, and longer hours of childcare.[11] This increased burden continued for the period of the school closures. Fathers and sons rarely helped out at home, although they tried to supplement lost family income. Families reported that boys generally had more mobility and free time to play, meet friends and study.

The eldest/elder daughters took on the most chores and shared a larger proportion of the household burden with their mothers. As one girl (high achiever, government school) put it to the interview team: "*I have one sister and two younger brothers and because I am the oldest, it's obvious that I have the most work-related responsibilities. I helped my mother with household work. Sometimes I washed the dishes or swept and mopped the house. If mother was doing the laundry, I would wring the clothes and hang them to dry. Sometimes, I would take my siblings out to play so that mother could get her work done.*" One family also indicated that while their son attended a private tutoring center, they did not send their daughters, as these were too far from the parental home.

Families showed high awareness of the negative effect of school closures and potential learning losses. Mothers, fathers and daughters were all concerned that the school closures would lead children to fall behind in school. They also were not sure about how much learning children would retain while at the parental home. Self-study was not seen as a good way to make up for lost instruction time in school, although siblings helped the younger children when they could. As one girl put it to us (low achiever, private school): *"Yes, I forgot a lot because I did not study anything. I could study some things that were easy. Additionally, my sisters and brother would explain the content that they understood. If they did not understand it then I would ask my friend for help and later revise it only if I understood it. However, I would abandon it if I did not understand it."*

Qualitative interviews provide clues for why gender-neutral messages may be more effective than gender-specific messages in a context like Pakistan. In father interviews, it seemed that parents did not differentiate in their explicit expectations between boys and girls and in general tended to support the child who required more assistance during school closure. Gender neutral messaging may appeal to such parents who did not find the difference between boys and girls to be salient in their thinking. Some mothers did mention that their daughters did not need tuition because they were intelligent and could study on their own or because their older siblings or parents could teach them.

The interviews also found important constraints on remote learning activities. Despite widespread news about opportunities for remote learning on television and online, families were simply often

[11] The results shown in Table 11 indicated that that the information campaign did not affect the extensive or intensive margins of girls' participation in household chores, despite a high burden of these chores on girls in the household (well over half of girls took up household chores, with the average time duration of 30 minutes per day). It should be noted that message content did not address girls' burden of household chores directly and rather emphasized the positive aspects of girls' (children's) education and the need to spend time on learning.

unaware of the channels through which they could access these opportunities. Social norms also played a role, however. Girls that we interviewed mentioned that their families did not allow them to watch television and were hesitant to let them use smartphones. Families associated watching television with watching soap operas or other shows that were tangential to learning. But even those who did watch these channels found that the non-interactive nature of television programming was not conducive to learning. As expressed by one of the interviewees (low achiever, private school): *"I used to like those TV programs while I watched them. They were useful. The problem was that we could not ask questions if we did not understand something. I had to call my mother and that topic would pass since we could not pause it."*

Together, these findings indicate that the school closures were a significant disruption to children's learning. Families were generally aware of these constraints but did not necessarily have the tools to remediate potential learning losses. Families did not generally distinguish strongly in these concerns between boys and girls, although their children did have very different lived experiences during the pandemic. The remote learning options available (particularly the television) were generally not regarded as an effective tool to continue studying. Text messages from the government were generally considered a positive contribution to children's education. Nevertheless, there were gendered communication barriers, with fathers controlling this flow of information to the rest of the household.

6. Discussion and Conclusion

Supporting families and children through information campaigns and providing remote learning opportunities has been a common way for governments to support learning through the pandemic. Overall, this literature has found important support for information campaigns and distance learning opportunities, although there also have been null effects. The intervention displayed here shows both some of the potential mechanisms and the limitations of this approach in a low-income setting like Pakistan. What is unique about this information campaign is that it is implemented by a government institution at scale, and it featured messages which sought to inform parents about the availability of education resources, to convince them to send their children back to school when they re-opened and to provide parents with actionable tips for pro-health and pro-learning behaviors during the pandemic, and keeping their children engaged with learning.

First, this paper found strong evidence that a text and voice messaging program can be effective in helping families to decide to re-enroll their children in school. Qualitative interviews confirmed that many fathers shared text messages with their families, showing that messages can reach entire households. The results provide strong evidence that messages can increase test scores after a long period of school closures. We observe that this effect is largest for students who are more likely to drop out, consistent with the idea the effect is concentrated among students who are on the margin of deciding to drop out. Finally, we have found that messages had a small positive effect among those using remote learning opportunities, while reducing the intensity of tutoring.

The results also suggest some potential pathways that may help us to understand how these effects came about. As discussed in the context, school closures were intermittent, and we only observe the large effect sizes on enrollment once schools were declared permanently open. While we are not formally able to identify pathways, the results suggest that learning was produced at least in part through re-enrolling in school, signaling the importance of schooling for learning. Similarly, the results

also suggest that gender-neutral messaging may be more effective in a context like Pakistan, where there may be important norms against girls' participation in learning. Qualitative interviews revealed that families generally did not find that the gender difference in learning losses were salient, perhaps making them more responsive to gender neutral messaging. This finding aligns with the idea that gender-neutral interventions may help girls as much or even more than gender-targeted interventions (Evans et al, 2022).

The effects of Treatment 3, which featured Academic Support messages, consisting of a series of practice problems and their answers to help children with their learning, were mixed. This intervention was carefully designed to help students with common problems they have in their regular textbooks and workbooks that are used in schools. While this program was implemented while schools were already re-opened, the results show that it raised re-enrollment rates (given that enrollment in Pakistan is a continuous process throughout the school year). However, the academic support messages did not have a statistically significant effect on test scores. This may lead us to conclude that academic support messages played a more important role in reminding parents about the importance of schooling and learning, rather than the content of the text messages themselves being a good tool to improve learning outcomes (Angrist et al, 2022).

These results also raise some important questions, however. First, we only observe the effect sizes on enrollment at a period when schools have permanently reopened. In early rounds of the survey, we observe a positive effect on re-enrollment, but the results are underpowered to assess whether this effect is statistically different from a null effect. The question this raises is whether credible official signals of their commitment to keep schools open play an important role in producing these effects. Given the observed effect sizes, it is possible that the content of the text and voice messages (which emphasized the permanent opening of schooling) may have been such a signal.

Though we do not detect any widespread or sizable effects on the time use of children as a result of the information campaign, it is still possible that the program led to changes in how parents talked about education at home and oriented their children toward learning. This is confirmed in the qualitative interviews, where mothers and fathers mentioned discussions about the messages in the household. We find that the point estimates of effects of the text and voice message program on Urdu and Math are fairly similar. We speculate that the fact that the effects on Urdu and Math are similar could be suggestive of the program affecting parental engagement with children's learning. Education studies commonly find that school inputs have larger effects on math than language test scores, which is consistent with schools accounting for the vast majority of learning in math but the family playing a larger role in language (Jackson, Rockoff, and Staiger 2014). If the test score improvements are only a result of increased school enrollment, we might have expected smaller effects on Urdu compared to Math.

The finding that the program had no significant effects on students' engagement with remote learning may be unsurprising for two reasons. First, schools had been permanently re-opened during endline data collection which may mean that school, doing homework, self-study or with a private tutor (although unobserved in round 3) could have been more important for children than the content of these practice homework problems. These reasons may also help to explain the lack of significant effect of the program on students' time spent on self-study. Second, the qualitative research indicated negative parental and children's perception with the use of television for educational purposes. These

included both social norms against television usage, and the low knowledge about and perceived quality of the remote learning available through television programming.

The qualitative research provided further possibilities to interpret these results. Interviewees were generally aware and concerned about learning losses among their children. The interviews indicated that most mothers and fathers had difficulties in providing academic support to their children's education, both due to income losses and the absence of schooling options. Girls also faced constraints as tutoring centers sometimes are physically distant from the parental home, and girls are not allowed to travel as far as boys. In this context, the treatment may have solidified the perceptions of the cost of education as a long-term investment, showing concern about the challenges faced by families, and providing timely information about returning to school and countering potential negative decisions on education.

Taken together, this research points to the importance of information in bringing about behavioral change, in particular in a relatively socially conservative environment such as Pakistan. The study leveraged COVID-19-related school closures to assess the efficacy of a low-cost behavioral nudge. Given the high penetration of mobile phones, text and voice messages potentially provide a scalable and effective solution not only in crisis situations, but also in non-emergency contexts to reach out to parents and address norms that may constrain children from attending school. In emergency situations, text and voice messages can support continuity of learning with practical information for parents and help reduce losses experienced by children in resource constrained settings.

Most importantly, the results of this study point to the continued importance of face-to-face schooling. In low-income settings like Pakistan, there often are no good at-scale alternatives for learning in school. This was certainly the case during the pandemic, where governments struggled to provide cost-effective solutions for children. As this paper shows, providing text and voice message support for families can be a feasible option for governments to encourage learning and re-enrollment. But the results also suggest the importance of schooling itself for learning. Perhaps the biggest lesson for policy makers faced with similar situations in the future will be to keep schools open as much as possible, within reasonable limits of public health impacts.

References

Ali, S. N., R. Bénabou (2020). "Image versus Information: Changing Societal Norms and Optimal Privacy." *American Economic Journal: Microeconomics* 12 (3): 116–64**.**

Al-Ubaydli, O., List, J. A., & Suskind, D. (2019). The Science of Using Science: Toward an Understanding of the Threats to Scaling Experiments. NBER Working Paper 25848.

Amin, S., Hossain, M., & Ainul, S. (2021). Learning loss among adolescent girls during the COVID-19 pandemic in rural Bangladesh. New York: Population Council.

Angrist, N., Bergman, P. & Matsheng, M. Experimental evidence on learning using low-tech when school is out. *Nature Human Behavior* **6**, 941–950 (2022).

Angrist N, Bergman P, Evans DK, Hares S, Jukes MCH, Letsomo T. (2020) Practical lessons for phone-based assessments of learning. *BMJ Global Health*.

Angrist, N., Ainomugisha, M., Bathena, S. P., Bergman, P., Crossley, C., Cullen, C., Letsomo, T., Matsheng, M., Panti, R. M., Sabarwal, S., & Sullivan, T. (2025). Building resilient education systems: Evidence from large-scale randomized trials in five countries (NBER Working Paper No. 31208). National Bureau of Economic Research. https://doi.org/10.3386/w31208

Andrabi, T., B. Daniels, and J. Das (2021). Human Capital Accumulation and Disasters: Evidence from the Pakistan Earthquake of 2005. The Journal of Human Resources,

Aslam, M. (2009) Education Gender Gaps in Pakistan: Is the Labor Market to Blame? Economic Development and Cultural Change, Vol 57(4) doi: 10.1086/598767

Azevedo, J. P., A. Hasan, D. Goldemberg, K. Geven, S. A. Iqbal (2021) Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates, *The World Bank Research Observer*, Vol 36(1), pp. 1–40

Banerjee, A., M. Alsan, E. Breza, A. G. Chandrasekhar, A. Chowdhury, E. Duflo, P. Goldsmith-Pinkham, B. A. Olken (2020). Messages on Covid-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients With Similar Effects on Non-Recipient Members ff Their Communities. NBER Working Paper 27496

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., & Walton, M. (2017). From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. Journal of Economic Perspectives, 31(4), 73–102.

Baron, J., M. Bend, N. Ejaz, J. D. Lee, I. Trako (2023). Breaking Barriers, Improving Futures Challenges and Solutions for Girls' Education in Pakistan. Washington DC: World Bank

Beam, E., S. Chaparala, S. Chaterji P. Mukherjee (2021) Impacts of the pandemic on vulnerable households in Bangladesh. IGC

Berlinski, Busso, Dinkelman & Martínez 2021; Reducing Parent-School Information Gaps and Improving Education Outcomes: Evidence from High-Frequency Text Messages. NBER Working Paper 28581

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

Betthäuser, B.A., Bach-Mortensen, A.M. & Engzell, P. A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nat Hum Behav 7, 375–385 (2023).

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (2020). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review* 110 (10): 2997–3029.

Crawfurd, L., Evans, D. K., Hares, S. & Sandefur, J. *Teaching and Testing by Phone in a Pandemic*. Working Paper No. 591 (Center for Global Development, 2021)

Dela Cruz, N. A., A. J. Adona, R. Molato-Gayares, A. Park (2025) Learning loss and recovery from the COVID-19 pandemic: A systematic review of evidence, International Journal of Educational Development, Vol. 115

David K. Evans, Fei Yuan, 2019. "What We Learn about Girls' Education from Interventions that Do Not Focus on Girls." CGD Working Paper 513. Washington, DC: Center for Global Development.

De Walque, D., C. Valente (2023). "Incentivizing School Attendance in the Presence of Parent-Child Information Frictions." *American Economic Journal: Economic Policy* 15 (3): 256–85.

Escueta, M., A. J. Nickow, P. Oreopoulos, V. Quan (2020) "Upgrading Education with Technology: Insights from Experimental Research." *Journal of Economic Literature* 58 (4): 897–996.

Flor, L. S., Friedman, J., Spencer, C. N., Cagney, J., Arrieta, A., Herbert, M. E., Stein, C., Mullany, E. C., Hon, J., Patwardhan, V., Barber, R. M., Collins, J. K., Hay, S. I., Lim, S. S., Lozano, R., Mokdad, A. H., Murray, C. J. L., Reiner, R. C. Jr., Sorensen, R. J. D., Pigott, D. M., Haakenstad, A., & Gakidou, E. (2022). Quantifying the effects of the COVID-19 pandemic on gender equality on health, social, and economic indicators: A comprehensive review of data from March, 2020, to September, 2021. *The Lancet*, Volume 399 (10344), pp. 2381-2397

Geven, K., Hasan, A. (2020) Learning Losses in Pakistan Due to Covid-19 School Closures: A Technical Note on Simulation Results. Washington DC: World Bank.

Hernandez-Agramonte, Juan Manuel and Namen, Olga and Naslund-Hadley, Emma and Biehl, Maria, Supporting Early Childhood Development Remotely: Experimental Evidence from SMS Messages. Journal of Development Economics, Vol. 166, doi:

Jacoby, H. G., Mansuri, G. (2011). Crossing Boundaries. Gender, Caste and Schooling in Rural Pakistan. World Bank Policy Research Working Paper 5710.

Jordan, K., K. Damani, C. Myers, A. Zhao. 2023. The Use of SMS and Other Mobile Phone-based Messaging to Support Education at Scale: A Synthesis of Recent Evidence. In Proceedings of the Tenth ACM Conference on Learning @ Scale,

Jordan, K. (2023) How can messaging apps, WhatsApp and SMS be used to support learning? A scoping review. Technology, Pedagogy and Education, Vol. 32(3)

Ome, A., A. Menendez (2022) Using SMS and parental outreach to improve early reading skills in Zambia, Education Economics, 30:4, 384-398,

Kamran, Iram, Tahira Parveen, Rehan M. Niazi, Maqsood Sadiq, Fatima Azeem, Emily EunYoung Cho, and Karen Austrian. 2022. "Gendered effects of COVID-19 school closures: Pakistan case study," GIRL Center Research Brief no. 12. New York: Population Council.

Kidman, R., Breton, E., Behrman, J., & Kohler, H. P. (2022). *Returning to school after COVID-19 closures: Who is missing in Malawi?* International Journal of Educational Development, Vol. 93 (Article 102645).

Moscoviz, L., D. K. Evans (2022) Learning Loss and Student Dropouts during the COVID-19 Pandemic: A Review of the Evidence Two Years after Schools Shut Down. Washington DC: Center for Global Development.

Pakistan Institute of Education (2023). Pakistan Education Statistics 2019-20. Islamabad: Ministry of Federal Education & Professional Training, Government of Pakistan.

Rodriguez-Segura, D. (2022) EdTech in Developing Countries: A Review of the Evidence, *The World Bank Research Observer*, Volume 37(2), pp. 171–203

Sabarwal, S., A. Y. Chang, N. Angrist, R. D'Souza (2023) Learning Losses and Dropouts. The Heavy Cost COVID-19 Imposed on School-Age Children. In Schady, N., A. Holla, S. Sabarwal, J. Silva, A. Y. Chang (eds.) How the COVID-19 Pandemic Eroded Human Capital and What to Do about It. Washington DC: World Bank.

Schueler, B., D. Rodriguez-Segura (2023) A cautionary tale of tutoring hard-to-reach students in Kenya. Journal of Research on Educational Effectiveness, Vol. 16(3)

Tembey, L., Baier, J., Ogolla, C., & Mohan, P. (2021). Understanding barriers to girls' access and use of EdTech in Kenya during Covid-19. EdTech Hub.

UNESCO (2010) Secondary Education Regional Information Base: Country Profile: Pakistan. Bangkok: UNESCO.

UNESCO (2021) When School Shut. Gendered impacts of COVID-19 school closures. Paris: UNESCO.

Zulaika, G., Bulbarelli, M., Nyothach, E., Van Eijk, A., Mason, L., Fwaya, E., Obor, D., Kwaro, D., Wang, D., Mehta, S. D., & Phillips-Howard, P. (2022). *Impact of COVID-19 lockdowns on adolescent pregnancy and school dropout among secondary schoolgirls in Kenya.* BMJ Global Health, Vol. 7(1)

Appendix

A.1. Selection probabilities

For the baseline analysis, Sample 1 and Sample 2 were treated as separate populations. This is because the names of the schools in the phone number data provided to the schools were cloned by the government for privacy reasons, and as a result, it is not possible to identify the schools in the original sampling frame. Sample 1 is representative of Punjab Province by virtue of the sampling frame being the SABER survey (which was designed to be provincially representative), while Sample 2 is representative of the schools from which the phone numbers were provided.

For Sample 1, the population is defined as girls who were in 4th grade in 2018 with observable test score data and phone numbers for parents. Those girls who, if selected into SABER would not have test score data or phone numbers (e.g.: because of non-response) are excluded from the population. This is a technical definition, but it is needed to precisely define our population and to calculate the sample weights, and its explicit definition prompts us to consider how our target population may differ from the full population. From this definition, to be selected into Sample 1, the girl would first need to be selected into the SABER survey and have observable test score and phone number data and then second be available and willing to respond to the baseline survey for this study. If $stu_{1,i,j}$ is a binary variable denoting whether a girl in the ith stratum and jth school was selected at baseline, then the probability of this girl being selected, $P\{stu_{1,i,j}\}$, is defined as

$$P\{stu_{1,i,j}\} = P\{sch_{0,i,j}\}P\{stu_{0,i,j}|sch_{0,i,j}\}P\{sch_{1,i}|sch_{0,i}\}P\{stu_{1,i,j}|sch_{0,i,j}, stu_{0,i,j,}\}$$

Where

$P\{sch_{0,i,j}\}$ is the probability of school j in stratum i being selected in the SABER survey,

$P\{stu_{0,i,j}|sch_{0,i,j}\}$ is the probability of a student in school j in stratum i being selected into the SABER survey conditional on her school being selected,

$P\{sch_{1,i}|sch_{0,i}\}$ is the probability of schools in stratum i being selected conditional on their selection into the SABER survey (school response rate), and

$P\{stu_{1,i,j}|sch_{0,i,j}, stu_{0,i,j,}\}$ is the probability of a student in school j in stratum i being selected in the baseline survey conditional on being selected into the SABER survey.

Schools were selected into the SABER survey using probability proportional to size (PPS) sampling defined by enrollment numbers. By definition of PPS sampling,

$$P\{sch_{0,i,j}\} = m_i \frac{enrl_{i,j}}{ENRL_i} P\{schresp_{0,i}\}$$

Where:

m_i denotes the number of schools selected at baseline in stratum i,

$enrl_{i,j}$ denotes the enrollment in school j in stratum i,

$ENRL_i$ denotes the sum of enrollments in schools in stratum i, and

$P\{schresp_{0,i}\}$ denotes the probability of schools in stratum i responding to the SABER survey.[12]

$P\{sch_{1,i}|sch_{0,i}\}$ was defined as the response rate of schools that were surveyed in SABER in the baseline survey.

Because the SABER survey sampled all girls within the target population described previously, $P\{stu_{0,i,j}|sch_{0,i,j}\} = 1$. Also, as described previously, response rates to the baseline survey were low. $P\{stu_{1,i,j}|sch_{0,i,j}, stu_{0,i,j,}\}$ was calculated as the proportion of students sampled in SABER in the target population who were successfully contacted at baseline in each school. In analysis of the baseline data, it was found that girls with higher test scores were more likely to be successfully contacted in the baseline survey. As a result, $P\{stu_{1,i,j}|sch_{0,i,j}, stu_{0,i,j,}\}$ was alternatively estimated by further conditioning on students' SABER test scores using a logit model with selection as the dependent variable, and test scores as well as school binary variables as the independent variables. These weights will be used in subsequent analysis of Sample 1.

For Sample 2, the selection probabilities were defined as

$$P\{stu_{1,i,j}\} = P\{sch_{1,i,j}\}P\{stu_{1,i,j}|sch_{1,i,j}\}$$

Where

$P\{sch_{1,i,j}\}$ is the probability of school j in stratum i being selected and

$P\{stu_{1,i,j}|sch_{1,i,j}\}$ is the probability of a student within each school being selected conditional on the school being selected. Because the government provided the phone numbers for Sample 2 schools and because it is not possible to identify the schools, Sample 2 is treated as being a separate population. As a result, $P\{sch_{1,i,j}\} = 1$ and $P\{stu_{1,i,j}|sch_{1,i,j}\}$ was estimated defined as the proportion of the number of phone numbers in each school that were successfully contacted. Sample weights were defined as the inverse of the selection probabilities.[13]

[12] Note that PPS was used in the selection of schools in the SABER survey; however, a subsequent round of sampling of schools was conducted also by PPS in some strata to increase the sample size. Numerical methods were used to calculate the selection probabilities of each school in these strata by repeating this double PPS sampling method 1,000 times and estimating the relationship between frequency of selection and school size to predict the sampling probability for a school with a given size.

[13] Inverse selection probability weighting is used in complex survey analysis to provide unbiased estimates of totals (e.g.: Horvitz & Thompson 1952). The definition of $P\{sch_{0,i,j}\}$ provides the intuition behind why oversampled strata receive lower weights: increasing the number of schools in a stratum, m_i, increases the selection probability and subsequently reduces the weights. The sum of the selection probabilities, ignoring the non-response adjustment, in a stratum is equivalent to the proportion of the population sampled.

Appendix Tables

Appendix Table A.1 Attrition to Analytical sample (OLS)

	(1)	(2)	(3)	(4)
Any Treatment	0.061			
	(0.035)			
Treatment 1 (Gender-specific)		0.079		
		(0.044)		
Treatment 2 (Gender-neutral)			0.003	
			(0.036)	
Treatment 3 (Academic Support)				0.004
				(0.004)
R^2	0.36	0.44	0.49	0.48
N	4,041	2,704	2,669	1,404

All models control for imbalanced variable household size at baseline
* p<0.1; ** p<0.05; *** p<0.01***
Results are weighted.
School Fixed Effects included.
Standard Errors are denoted in parenthesis

Appendix Table A.2: Balance Table: Any Treatment vs. Control

	Control		**Any Treatment**		**Pairwise t-test**	
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	Mean difference
Age <13 (Tercile – 1)	721	0.544	1377	0.528	2098	0.016
	520	(0.044)	704	(0.032)	824	
Age 13 (Tercile – 2)	721	0.208	1377	0.247	2098	-0.039
	520	(0.031)	704	(0.026)	824	

Age >13 (Tercile – 2)	721	0.248	1377	0.224	2098	0.024
	520	(0.040)	704	(0.032)	824	
Grade enrolled at baseline	690	5.713	1334	5.770	2024	-0.057
	503	(0.057)	689	(0.048)	806	
Rural School	721	0.673	1377	0.584	2098	0.089
	520	(0.041)	704	(0.038)	824	
Asset Tercile - 1	721	0.333	1377	0.259	2098	0.075
	520	(0.042)	704	(0.022)	824	
Asset Tercile - 2	721	0.317	1377	0.309	2098	0.008
	520	(0.042)	704	(0.028)	824	
Asset Tercile - 3	721	0.347	1377	0.431	2098	-0.084
	520	(0.043)	704	(0.034)	824	
Total Adults in HH	721	3.561	1377	3.585	2098	-0.023
	520	(0.147)	704	(0.120)	824	
Household size	721	3.118	1377	2.820	2098	**0.298****
	520	(0.105)	704	(0.075)	824	
Father Not Educated	721	0.326	1377	0.223	2098	0.103
	520	(0.048)	704	(0.022)	824	
Father's Education - Less than Grade 10	721	0.354	1377	0.416	2098	-0.062
	520	(0.038)	704	(0.032)	824	
Father's Education - Grade 10 and above	721	0.320	1377	0.361	2098	-0.041
	520	(0.041)	704	(0.036)	824	
Mother Not Educated	721	0.563	1377	0.430	2098	0.134
	520	(0.043)	704	(0.031)	824	
Mother's Education - Less than Grade 10	721	0.282	1377	0.357	2098	-0.075
	520	(0.033)	704	(0.033)	824	
Mother's Education - Grade 10 and above	721	0.155	1377	0.213	2098	**-0.058***
	520	(0.033)	704	(0.031)	824	

Appendix Table A.3 - Balance Tests Treatment 1 vs. Treatment 2 vs. Control

	Control		T1		T2		Pairwise t-tests					
							T1 vs. Control		T2 vs. Control		T1 vs T2	
	N. Clusters	Mean (SE)	N. Clusters	Mean (SE)	N. Clusters	Mean (SE)	N. Clusters	Mean Difference	N. Clusters	Mean Difference	N. Clusters	Mean Difference
Age <13 (Tercile – 1)	721	0.544	708	0.574	669	0.478	1429	-0.030	1390	0.066	1377	0.096
	520	(0.044)	526	(0.041)	504	(0.047)	729		713		704	
Age 13 (Tercile – 2)	721	0.208	708	0.263	669	0.231	1429	-0.055	1390	-0.023	1377	0.032
	520	(0.031)	526	(0.040)	504	(0.032)	729		713		704	
Age >13 (Tercile – 2)	721	0.248	708	0.163	669	0.291	1429	0.085	1390	-0.043	1377	-0.128
	520	(0.040)	526	(0.022)	504	(0.056)	729		713		704	
Grade enrolled at baseline	690	5.713	681	5.753	653	5.788	1371	-0.040	1343	-0.076	1334	-0.036
	503	(0.057)	510	(0.053)	495	(0.075)	706		699		689	
Rural School	721	0.673	708	0.597	669	0.571	1429	0.076	1390	0.103	1377	0.026
	520	(0.041)	526	(0.046)	504	(0.052)	729		713		704	
Asset Tercile - 1	721	0.333	708	0.249	669	0.269	1429	0.084	1390	0.064	1377	-0.021
	520	(0.042)	526	(0.029)	504	(0.035)	729		713		704	
Asset Tercile - 2	721	0.317	708	0.345	669	0.270	1429	-0.029	1390	0.047	1377	0.076
	520	(0.042)	526	(0.043)	504	(0.036)	729		713		704	
Asset Tercile - 3	721	0.347	708	0.405	669	0.459	1429	-0.058	1390	-0.112	1377	-0.054
	520	(0.043)	526	(0.043)	504	(0.051)	729		713		704	
Number of Adults in the Household	721	3.561	708	3.685	669	3.476	1429	-0.124	1390	0.086	1377	0.209
	520	(0.147)	526	(0.169)	504	(0.174)	729		713		704	
Number of Children in the Household	721	3.118	708	2.765	669	2.880	1429	**0.353****	1390	0.238	1377	-0.115
	520	(0.105)	526	(0.112)	504	(0.080)	729		713		704	
Father Not Educated	721	0.326	708	0.232	669	0.214	1429	0.094	1390	0.112	1377	0.018
	520	(0.048)	526	(0.031)	504	(0.028)	729		713		704	
Father's Education - Less than Grade 10	721	0.354	708	0.458	669	0.370	1429	-0.104	1390	-0.016	1377	0.088
	520	(0.038)	526	(0.042)	504	(0.041)	729		713		704	
Father's Education - Grade 10 and above	721	0.320	708	0.310	669	0.416	1429	0.010	1390	-0.096	1377	-0.106
	520	(0.041)	526	(0.039)	504	(0.053)	729		713		704	
Mother Not Educated	721	0.563	708	0.418	669	0.443	1429	0.146	1390	0.121	1377	-0.025
	520	(0.043)	526	(0.040)	504	(0.045)	729		713		704	

Mother's Education - Less than Grade 10	721	0.282	708	0.383	669	0.329	1429	-0.101	1390	-0.047	1377	0.054
	520	(0.033)	526	(0.043)	504	(0.053)	729		713		704	
Mother's Education - Grade 10 and above	721	0.155	708	0.199	669	0.229	1429	-0.044	1390	-0.074	1377	-0.029
	520	(0.033)	526	(0.038)	504	(0.042)	729		713		704	

Notes:

***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Appendix Table A.4: Balance Table: Treatment 3 (Academic support) vs. Control

	Control		**Treatment (T3 – Academic support)**		Pairwise t-test T3 vs. Control	
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	Mean difference
Age <13 (Tercile – 1)	721	0.544	680	0.561	1401	-0.017
	520	(0.044)	489	(0.043)	708	
Age 13 (Tercile – 2)	721	0.208	680	0.239	1401	-0.031
	520	(0.031)	489	(0.040)	708	
Age >13 (Tercile – 2)	721	0.248	680	0.200	1401	0.048
	520	(0.040)	489	(0.027)	708	
Grade enrolled at baseline	690	5.713	655	5.803	1345	-0.090
	503	(0.057)	476	(0.062)	689	
Rural School	721	0.673	680	0.574	1401	0.100
	520	(0.041)	489	(0.048)	708	
Asset Tercile - 1	721	0.333	680	0.263	1401	0.070
	520	(0.042)	489	(0.031)	708	
Asset Tercile - 2	721	0.317	680	0.313	1401	0.003
	520	(0.042)	489	(0.042)	708	
Asset Tercile - 3	721	0.347	680	0.421	1401	-0.074
	520	(0.043)	489	(0.045)	708	
Number of Adults in the Household	721	3.561	680	3.612	1401	-0.051
	520	(0.147)	489	(0.182)	708	

Number of Children in the Household	721	3.118	680	2.905	1401	0.213
	520	(0.105)	489	(0.118)	708	
Father Not Educated	721	0.326	680	0.245	1401	0.081
	520	(0.048)	489	(0.034)	708	
Father's Education - Less than Grade 10	721	0.354	680	0.409	1401	-0.055
	520	(0.038)	489	(0.045)	708	
Father's Education - Grade 10 and above	721	0.320	680	0.346	1401	-0.026
	520	(0.041)	489	(0.044)	708	
Mother Not Educated	721	0.563	680	0.429	1401	0.135
	520	(0.043)	489	(0.040)	708	
Mother's Education - Less than Grade 10	721	0.282	680	0.382	1401	-0.101
	520	(0.033)	489	(0.044)	708	
Mother's Education - Grade 10 and above	721	0.155	680	0.189	1401	-0.034
	520	(0.033)	489	(0.037)	708	

Appendix Table A.5. Effect of Any Treatment on (1) Re-enrollment, (2) Urdu Scores, (3) Math Scores at Endline, using all controls from balance table

	Re-enrollment	Urdu Test Scores	Math Test Scores
Any Treatment	0.057*	0.209**	0.209**
	(0.029)	(0.091)	(0.106)
Age Tercile - 2	0.005	0.013	0.126
	(0.030)	(0.090)	(0.106)
Age Tercile - 3	-0.036	-0.040	0.028
	(0.033)	(0.139)	(0.147)
Rural School	-0.062	0.038	0.220
	(0.048)	(0.132)	(0.133)
Asset Tercile - 2	0.028	0.007	-0.042
	(0.028)	(0.095)	(0.113)
Asset Tercile - 3	0.024	0.040	-0.034
	(0.033)	(0.105)	(0.141)
Missing data on asset score	0.031	0.314	0.775*
	(0.050)	(0.267)	(0.400)
Total Adults in Household	-0.001	-0.013	-0.022
	(0.005)	(0.016)	(0.022)
Household size	0.017	0.098***	0.087**
	(0.012)	(0.038)	(0.043)
Father's Education - Less than Grade 10	0.006	0.285***	0.170
	(0.037)	(0.106)	(0.120)
Father's Education - Grade 10 and above	0.019	0.303**	0.304**
	(0.039)	(0.130)	(0.137)
Mother's Education - Less than Grade 10	0.032	0.139	0.156
	(0.025)	(0.100)	(0.111)
Mother's Education - Grade 10 and above	0.015	0.160	0.267**
	(0.036)	(0.123)	(0.132)
Child was enrolled in Grade 6 at baseline	0.019	-0.001	0.142
	(0.040)	(0.140)	(0.139)
Child was enrolled in Grade 7 at baseline	0.002	-0.143	0.124

	(0.061)	(0.162)	(0.178)
Missing data on grade at baseline	-0.141	-0.224	-0.055
	(0.088)	(0.251)	(0.295)
Mean value of dependent variable in the control group	*0.926*		
R^2	0.55	0.65	0.64
N	2,098	2,098	2,098

Results are weighted. Standard Errors are denoted in parenthesis. Results include school fixed effects. * p<0.1; ** p<0.05; *** p<0.01*