Evaluation of a Parental Engagement Program

Endline report on impacts of a parental engagement SMS intervention

IDinsight

Evaluation of a Parental Engagement Program in Lilongwe Rural Endline report on student learning outcomes following a parental engagement SMS intervention

April 2023

Authors
Arthur Chibwana: **Arthur.Chibwana@IDinsight.org**
Jack Thunde: **Jack.thunde@IDinsight.org**
Mtise Mwanza: **Mtise.mwanza@IDinsight.org**
Valentina Brailovskaya: **Valentina.brailovskaya@IDinsight.org**

About IDinsight
IDinsight uses data and evidence to help leaders combat poverty worldwide. Our collaborations deploy a large analytical toolkit to help clients design better policies, rigorously test what works, and use evidence to implement effectively at scale. We place special emphasis on using the right tool for the right question, and tailor our rigorous methods to the real-world constraints of decision-makers.

IDinsight works with governments, foundations, NGOs, multilaterals and businesses across Africa and Asia. We work in all major sectors including health, education, agriculture, governance, digital ID, financial access, and sanitation.

We have offices in Dakar, Lusaka, Manila, Nairobi, New Delhi, Rabat, and Remote. Visit **www.IDinsight.org** and follow on Twitter **@IDinsight** to learn more.

Table of Contents

Abstract

We present results of a randomised control trial aimed at estimating causal impacts of SMS nudges aimed to improve parental engagement with their children. The study took place in a single TA in Lilongwe Rural, Malawi among 2,095 active GiveDirectly recipients. Half of GiveDirectly cash transfer recipients were randomised to receive messages on their phones with information on the importance of education and tips on how to engage with their children. The control group only received GiveDirectly transfers. We find no impacts across all pre-specified indicators. Contrary to experimental findings, we find overwhelming interest in the messages among recipients of the intervention and large self-reported impacts. We find substantial spillovers (much larger than anticipated) and about half of the treatment group reports sharing the messages with other people in their community. Roughly about the same percentage of the control group reports either receiving messages or discussing messages with other people. This limits our ability to conclusively state whether the intervention is impactful on the main indicators of interest– parental engagement and child school attendance.

Given the inconclusive results, we cannot recommend scaling up the intervention with certainty. However, a few considerations such as relative cost of the intervention and heightened enthusiasm among recipients, could help GiveDirectly and interested funders make decisions. If the intervention is scaled, we advise further tailoring messages to the local context, conducting periodic descriptive studies to measure interest in the intervention by recipients, and systematically tracking important indicators (such as attendance rates) against pre-selected targets.

1. Introduction

Increased parental engagement in their children's lives has been shown to improve economic, psychological, and socio-emotional outcomes many years later (Gertler et al. 2021, Walker et al. 2021). However, parents, especially those in developing countries, may have competing demands for their attention, including thinking about next earning opportunities for casual labourers, planning for the agricultural seasons and tending to crops, and accumulating savings to insure against health shocks among others. (Munje & Mncube 2018). Given multiple demands, it might be challenging to remember to dedicate adequate time to build stronger bonds with children, and interventions that effectively teach parents on what to do as well as increase salience of engagement with children may be necessitated.

In this research, we experimentally study how simple SMS reminders sent twice a week with tips about effective parenting affect parent-child interaction and, ultimately, schooling outcomes. The intervention aims to improve information access to proven parental strategies, which may not be available to parents in rural areas. This includes the importance of school and nudging parents to build stronger relationships with their children.

The SMS content has been developed by Movva— an NGO that combines behavioural economics and artificial intelligence to improve educational indicators by utilising nudges containing general information about the value of parental engagement in child education. Previous literature shows promising impacts of such reminders on similar outcomes across contexts. For example, in the United States, York et al. 2019 shows that similar SMS reminders increased parental engagement by between 0.15-0.29 standard deviations (s.d.) after 8 months of program exposure. Angrist et al. 2021 demonstrate that SMS messages with learning activities improve learning by 0.12 SD in Botswana.

This SMS intervention has been tested in Brazil, Côte D'Ivoire, and Ghana[1]. In Brazil a year-long intervention has shown promising effects on educational outcomes. Bettinger et al. 2020 find that while messages with child-specific information had an impact only in the subject of interest (e.g. Math/language), general information about the importance of education (referred to as salience) was able to enhance educational outcomes across different subjects. The salience messages led to a significant increase in outcomes: 2.1 percentage points for attendance, 0.09 s.d. for Math GPA and standardised test scores, 3.2 percentage points for grade promotion rates and student effort by 0.019 s.d., when compared to the pure control group. The authors attribute these findings to the messages leading parents to ask their children more about school than those in the control group (engagement between parents and children increased between 0.09 and 0.12 standard deviations, compared to the control group). Bettinger et al. 2021 further

---

[1] Results for Cote D'Ivoire are yet to be available to the public

demonstrated that the positive impact of messages on attendance and GPA not only persisted but increased throughout the school year.

The use of SMS-based interventions has been observed to go beyond just academic improvements, with measures to prevent school dropouts and improve socio-emotional well-being. Lichand & Christen 2021 found that during the Covid-19 pandemic, SMS interventions encouraging parents to support their children during challenging times were effective in reducing dropout risks by 26% over the course of the year. Framing of the messages was identified as crucial, as messages that highlighted the benefits of graduating high school outperformed those that focused on the negative consequences of dropping out. Another study aimed to increase engagement with remote learning by sending targeted messages on socio-emotional skills to either students or caregivers (Lichand et al. 2022). The study demonstrated that treatment students not only outperformed their peers, but also experienced lower learning losses of 7.5% in mathematics and 24% in Portuguese.

This experiment took place in a single traditional authority in Lilongwe Rural in Malawi, with the total population of 140,536[2]. The study participants were GiveDirectly recipients in 32 villages who were scheduled to receive unconditional cash transfers of $621 USD over a period of 12 months. The research team at IDinsight was brought on to evaluate the causal impacts of education-related SMS nudges on GiveDirectly recipients to inform scale-up decisions of the intervention to the broader GiveDirectly population in Malawi. The research sample was recruited during the GiveDirectly registration process and consisted of households which reported having at least one child of school-going age. Upon recruitment into the study, the households were randomised into treatment and control groups by the IDinsight team. Since the recipient recruitment by GiveDirectly was done in three waves spanning 8 months, the treatment group received an intervention varying between 23 and 9 weeks, depending on when the individuals were registered by GiveDirectly.

For the purposes of the study and in line with the scale-up strategy, only 1 guardian from each household was enrolled in the study. Upon enrollment into the study, the treatment group started receiving SMS nudges with three main goals: (1) to inform guardians that engaging with children is important (2) provide concrete strategies on how to engage with children (3) encourage guardians to pay attention to children's schooling outcomes. The messages were delivered in the local language– Chichewa.

The baseline survey took place during GiveDirectly's first transfer follow-up interaction . The endline took place after the guardians had been exposed to between 3 and 23 weeks of the intervention and consisted of phone interviews conducted with guardians and their children, as well as in-person visits to schools to check attendance registers and progress reports.

The main outcomes of interest include child's school attendance and parent-child engagement.

---

[2] https://knoema.com/MWDS2018/demographic-statistics-of-malawi?tsId=1186430

First, **we note that the receipt of messages in the treatment group was high** and 90% of the treatment group reports receiving at least 1 Movva message. We know with certainty from Movva records, that about 95% treatment sample received messages (i.e. Movva received the delivery confirmation). The remainder of 5% of recipients who do not report receiving the messages likely didn't pay attention to them. Comprehension of the message was high with 93% of the treatment respondents reporting understanding what the messages were about, which is confirmed by the qualitative follow-up where respondents correctly identified content of the messages.

Second, **we find that there was overwhelming interest in receiving the messages.** All treatment respondents who acknowledged receiving messages report that they would be interested in receiving these types of messages in the future and 97% of them found messages to be either very useful or extremely useful.

Third, by investigating spillover effects from the treatment to the control group, **we find significant sharing of information between households**: about half of treatment respondents report either sending messages or discussing messages with members outside of their household. About the same fraction of the control group report that the content of the messages was shared with them either through direct receipt of messages or discussion with other community members. While sharing of messages was common, we do not believe that control respondents received the majority of the content. Conditional on sharing messages at all, an average treatment recipient reports sharing about 5 messages (range between 1 and 50) with about 5 people (range between 1 and 100). This is a surprising result, In fact, research aiming to utilise social networks for information dissemination do not reach this level of sharing. For example, in an agricultural setting, researchers document that less than 20% of households share information about agricultural practices. We attribute this phenomenon to a few factors. The geographical area of the intervention is quite small and the 32 villages of interest cover about 7,556 households in total. Since the randomization was done at the household level, it's possible that households within the same social networks were in both treatment and control groups. Since GiveDirectly is piloting the cash transfer program in a small area, it was impossible to expand the study area to more individuals and randomise at a higher level to fully prevent spillover effects. The second factor is that the sample consisted of GiveDirectly recipients who were actively receiving transfers and possibly sharing information about other benefits they got. Our third possible explanation is that the intervention generated a lot of enthusiasm among recipients, and they wanted to share what they have learned within their social networks. The latter is well-supported by the overwhelming interest in the messages expressed by the treatment group. However, we caution against making long term predictions based on this short term study. It's possible that initial enthusiasm for the messages may assuage with time.

Fourth, **we find no impacts on any indicators of interest** (either in attendance or parental engagement behaviours). In the presence of such significant spillover effects, it's challenging to conclude whether the intervention actually affected the behaviours of interest.

We also can't rule out that the effects are too small to detect with the given sample size. While the magnitude of impact is in the right direction, the effects on attendance and parental behaviour index are between 0.0 and 0.05 standard deviations[3] (against the control group), respectively. If this is the true effect size of the intervention, we would have needed a sample that is nearly 5x as large to detect significant impacts.

Furthermore, while the messages aim to directly induce some of the measured behaviours (e.g. talking with children), the primary outcome - attendance - was only explicitly encouraged in the final set of messages. It is also possible that the exposure to messages may have shifted children's behaviours on the intensive margin (e.g. more interested in school, working harder in school) and not on the extensive margin such as whether or not a child attended school on a randomly selected day. However, there is limited support for that as we do not find impacts on a child's intrinsic motivation to go to school or do homework.

Lastly, when we asked children why they didn't go to school, the main reason for non-attendance was sickness, which will not be changed by the intervention. The second main reason was not having a clean uniform or school supplies, which results in automatic dismissal from school in rural Malawi. Since this barrier was not explicitly mentioned in text messages for parents to think about, **it is possible that the SMS reminders didn't address the most pressing barrier to attendance.** It is also possible that the attendance rate is already high (at 90% in the control group at endline) and it's challenging to move that further.

We conclude with the following recommendations for the scaleup plans.

Given such high enthusiasm for the intervention and relatively low cost of implementation, **GiveDirectly should still consider scaling the intervention to other areas. We recommend that the scale-up be followed by an M&E plan and empirical measurement of key indicators.** The M&E plan could outline indicators that define the success of the intervention as well as levels of indicators which inform decisions for whether intervention should be continued. We suggest complementing the efforts with periodic follow-up surveys with recipients to assess their interest in messages and regularly measure important indicators. Success of the intervention implementation may be measured via pre-specified targets for the indicators of interest. For example, one of the indicators could measure their engagement by asking a random set of participants to recall the content of the last message. If more than 50% identify it correctly, then the intervention is continued, if the value is less than 50%, it could be considered obsolete. While we saw tremendous interest in the messages during this evaluation, it is unclear whether it will be sustained beyond the 5 months measurement window. Once interest in the message assuages, the intervention can be discontinued.

Given that the messages didn't address the primary reasons for non-attendance in this

---

[3] These effects are pretty small compared to other studies in education. Evans et al. 2022 synthesises the education literature and finds that the median effect size on learning is about 0.1 standard deviations compared to the control group. Given that attendance precedes learning gains, it's likely that impacts on attendances should be higher to achieve 0.1 standard deviation gains.

context, we recommend **conducting qualitative interviews with a subset of the target population to gain a contextual understanding of their challenges in order to better tailor the intervention to their specific needs**. The same interviews can be conducted on a yearly basis and contents of the messages can be tailored to the unique contextual factors.

The report is structured as follows: First, we describe evaluation methods employed in this study for participant selection, randomization, data collection and analysis, including the analytical model specification. Second, we dive into the empirical results of the evaluation touching on (i) balance across the groups, (ii) takeup of the messages (iii) the impact of the intervention on the primary and secondary research outcomes- attendance and non-school activities/behaviours, and (iv) limitations affecting our evaluation. Lastly, we propose some next steps GiveDirectly could take if they choose to scale-up the program.

2. Study Design

2.1 Sample selection and randomisation

2.1.1 Sample selection

This GiveDirectly cash transfer pilot program took place in 32 of 209 villages in a TA in rural Lilongwe. Field officers hired by GiveDirectly first conducted a census in 27 of 32 villages to map out every poor household. The 27 villages are under 6 Group Village Headmen (GVH). Between 21st March and 23rd December, 2022, GiveDirectly registered every eligible adult in all 32 villages (18 years and older) with a national ID in each household to receive GiveDirectly transfers[4]. Overall, 83-96% of households across the 32 villages were enrolled onto the cash transfer program. There were two processes that GiveDirectly followed to enrol recipients: (1) In-person enrolment and (2) remote self-enrolment through tele-targeting. For targeting, some villages were saturated with all adult residents above 18 being enrolled, and in some villages eligibility was determined based on a Proxy Means Test (PMT). For teletargetting, an algorithm was created using call data records from Airtel and a ground truth consumption survey to analyse phone usage patterns of individuals that were pegged to specific cell towers in 5 of the 32 villages, thereby identifying people living in poverty that could be targeted. Eligible households were enrolled on their phones via Unstructured Supplementary Service Data (USSD). In-person PMT surveys were then conducted in a teletargetting mop-up exercise to capture eligible individuals from 7 villages, who were unable to self enrol through USSD due to not having phones. Table 1 highlights some recruitment details.

Table 1: Recruitment means and timelines

Type of Recruitment	In-Person Recruitment				Teletargeting		Teletargeting Mop Up	
	Saturation		PMT		USSD		PMT	
	Dates	Days	Dates	Days	Dates	Days	Dates	Days
Time of Recruitment	21 Mar - 7 Apr 5 May - 11 Jun 8 Dec - 23 Dec	14 27 12	9 Jun - 1 Aug	38	11 Oct - 13 Oct	3	24 Oct - 16 Dec	40

---

[4] Recipients were allowed to have a trustee if they did not have an ID, however they did not receive their transfers unless they got an ID.

2.2 Baseline data collection

GiveDirectly collected baseline data from all participants in the 32 villages on a rolling basis. This was done during follow-up surveys administered remotely through a call centre and in-person via follow-up visits.

During the registration survey, the recipient received a phone, a SIM card and a flyer, which provided concise information on project information and mobile money usage[5].

The baseline survey included themes on guardian status (number of guardians in a household and demographic characteristics); child school enrolment and attendance status; child engagement in house activities; participant's willingness to receive Movva SMSs and consent to have their contacts shared with IDinsight and Movva and to be contacted to be part of the research study.

2.1.2 Exclusion criteria and randomisation

GiveDirectly lists of recipients and their children were given to IDinsight for randomization. Listed recipients had at least 1 child of school-going age and had agreed for their phone number to be shared with partner organisations (IDinsight and Movva). Prior to randomization we excluded the following households/children:

1. recipients who did not agree for their contact information to be shared with research partners, to be a part of an SMS research study or who did not have school-age children
2. Children that were 18 years old (1,101 across all three waves)
3. Children who were expected to be hard to reach either during phone surveys or in-person school visits.[6]

Before randomization at the household level, we randomly selected 1 adult guardian of each household to participate in the study and selected children reported by the randomly selected household member[7].

We randomised all households provided by GiveDirectly into treatment and control groups using Stata v17. Randomisation was stratified on child and parental characteristics[8]. Key stratification variables included: 1) age group of a randomly selected child 2) sex of the randomly selected child 3) distance to school (top/bottom 50th percentile), and 4) whether a parent reported ever talking to a teacher (proxy for parental engagement). Randomization strategies were very similar across waves, with an exception of redefining some of the categorical variables due to the fact that narrow

---

[5] Phones and SIM cards were only given to recipients who didn't have or chose to receive new ones from GD

[6] We excluded children who attended schools outside of the TA (wave 1=114; wave 2=16; wave 3=273) and children who attended schools with low representation in the sample (wave 1=24, wave 2=11, wave 3=18).

[7] At endline, we discovered discrepancies between the baseline records and children endline parental reports. In these cases, we enrolled children who were not originally registered by GiveDirectly if the children who were listed at baseline were not available to participate.

[8] The final sample contained 62 district strata, with an average size of 37 household (minimum 10, maximum 123)

definitions resulted in small stratum sizes. Our balance tables are in **Table 2 in Section 3.1**.

The final sample consisted of 2333 recipients in unique households randomised in 3 waves: 1513 in July 2022 (Wave 1), 382 in October 2022 (wave 2), 438 in November (wave 3). The full project timeline is depicted in **Appendix C**.

During data collection stages, we further excluded children that had not lived together with the guardian registered for the study in the past 6 months since the children weren't fully exposed to the intervention[9].

2.3 Endline data collection

2.3.1 Phone surveys

IDinsight hired a team of 34 enumerators and trained them on the objectives of the study, how to ethically conduct interviews, how to administer the survey and how to collect data using SurveyCTO. Guidance on how to download SurveyCTO onto their smartphones or tablets provided by IDinsigt was provided. Special emphasis was placed on child protection training to ensure children were protected during and after the call. This included materials on how to build rapport with both guardians and children by asking a series of questions about their education and aspirations among others. The full child protection protocol is listed in **Appendix D**.

Endline data was collected in two Rounds[10] to ensure we were able to complete data collection within agreed upon evaluation timelines.

Round 1: November 29 to December 3- occurred right before the start of term 1 exam period. We avoided surveying during the exams and the last week of the year because attendance is known to be higher during exams and lower right after exams are over.

Round 2: January 23[11] to February 6- took place after the intervention was fully concluded.

The endline survey contained four main sections: (i) Parental section - focused on respondent demographics such as education level, literacy, age and gender, aspirations for their children school outcomes for child of interest, engagement with child's school life, and parental child discipline practices; (ii) Child section - focused on similar

---

[9] We randomised the order of children within a HH; listed them on a roster and had enumerators ask parents, in order of randomization, if a selected child had lived with them for ≥ 5 days in a typical week in the past 6 months. Child was considered eligible for the study if they parent reported living with them in the past 6 months

[10] We had initially hoped to collect all data in December towards the end of the academic year. However, the ministry of education adjusted the 2022 and 2023 learning periods to normalise the school calendar such that it runs from October to September.

[11] Phone surveys were delayed by 2 weeks due to delayed school openings attributed to a cholera outbreak

school/education related themes, including engagement with guardian, and school motivation; (iii) Validation questions to calculate the correlation between parent/child reports on general engagement with child and receipt/understanding of SMS message; (iv) Reflection section, delivered to the treatment group, interrogating guardian's overall view of Movva SMSs, how useful they found SMSs to be, whether they experienced any adverse or positive effects, and if they would be interested in receiving similar messaging in the future.

All respondents were randomised into the date of the call and enumerator to ensure equal representation across waves and treatment groups on all dates of surveying.

During Round 1, enumerators completed 39% of the 2,333 research participants distributed across the 3 waves. The rest of the participants were surveyed in Round 2. Phone interviews with both guardians and children lasted about 30-45 minutes in total.

Enumerators pre-scheduled interviews with guardians to (i) check whether the randomly selected child or any other children lived with the guardian/parent in the past 6 months, and (ii) find a time in which both they and their child would be available at the same time. Once calls had been scheduled, an SMS reminder was sent to the participants the night before and the morning of the call. If the participants were unavailable at the agreed upon time, calls were rescheduled up to 7 times. Given the fact that these phone numbers were recently registered, we expected a high response rate and all households which were not reached after the original attempt were tried 7 more times. Overall, 6.6% of the sample was unreachable after 7 attempts, 2.7% were ineligible[12] and 0.3% refused to participate (**Table 11**).

Guardians were asked to consent to participate in the study for themselves and on behalf of their child. After enumerators had established rapport with the guardian during the interview and conducted the guardian's portion of the survey, they asked the guardian to hand the phone over to the child and began the child interview. Before interviewing the child, we obtained child assent.

During data collection, field managers conducted spot checks of enumerator calls to ensure quality. The IDinsight team also ran high-frequency checks (HFCs) and conducted daily debriefs with enumerators to address any issues in data quality. Back checks were also conducted with selected respondents where the research team had follow up questions.

2.3.2 In-person attendance data collection

Shortly after the end of the intervention in January 2023, in-person visits to collect data on attendance and test scores were conducted across 23 schools over 8 days with a team of 11 enumerators.

Head/teachers were asked to avail enumerators with ledgers of student attendance for

---

[12] Participants were deemed ineligible if no children had lived with the guardian for at least 5 days in a week during the 6 months of the intervention, if the guardian/parent was deceased or if the children of interest were deaf.

all terms of interest as well as class progress books for end of 2022/23 first term for collection of test scores. The team aimed to collect 1 attendance data point at baseline (July during the 2022 academic year) and 4 attendance points at endline (during the 2022-2023 academic year). For the endline, two attendance points were for November, close to the end of the intervention and two attendance points were for January, after completion of the intervention.

To maximise chances of finding the children, enumerators searched for student records at schools provided by both parent and child and where possible and looked for the children in registers 2 classes up and 2 classes down to what the parent or child reported. To preserve the privacy of participants, we did not disclose the names of children enrolled in the study to school personnel.

2.4 Specification

The average treatment effects (ATE) of Movva's SMS intervention on various outcomes will be estimated using the following Ordinary Least Squares model:

$$Y_i = \beta_o + \beta_1 T_i + \beta_2 Y_o + \beta_3 X_i + \delta_1 M_i + \alpha_i + \gamma_t + \varepsilon_i$$

Where:
- Y_i denotes the outcome variable for child i
- T_i denotes the treatment status of household j (1 for Treatment Group; 0 for Control Group)
- Y_o denotes the outcome variable for child i at baseline, where applicable
- X_i baseline characteristics - These are included to improve precision of the impact estimate.
    - HH-level
        - Household farms
        - Household does casual labour
        - Number of school-age children in the household
    - Age of parent
    - Gender of the parent
    - Whether parent is married or not
    - Whether parent ever visits school or talks to the teacher (engagement proxy)
    - Child-level
        - Age of child
        - Ever repeated grade
        - Days missed in the past 1 week
        - Total hours of chores

- When X_i is missing for an observation, we include an indicator term for missingness M_{io} and replace X_i with its mean

- β_1 is the estimated treatment effect of Movva's SMS intervention compared to control

- α_i strata fixed effects

- γ_t calendar week fixed effects instead.

- ε_i error term

The above equation will be estimated using OLS, with Huber-White robust standard errors. We registered our study design and pre-analysis plan on the American Economic Association registry for RCTs ahead of endline data collection.[13] As per the pre-analysis plan, we do not apply multiple hypotheses testing correction to the list of primary and secondary pre-specified outcomes, however, we do apply false discovery rate (FDR) adjustment to all other estimates following Benjamini et al. (2006) and reported adjusted q-values instead of the original p-values.

---

[13]RCT ID: AEARCTR-0009879, socialscienceregistry.org/trials/9879

3. Results

3.1 Balance

Table 2 reports average values for control and treatment groups for baseline household, parent and child characteristics in columns 1 and 2, respectively. The difference in means is reported in column 3. The stars next to the estimates signify whether the difference between the groups is statistically significant. Number of observations are reported in column 4. Overall, we conclude that the reached sample of respondents is well-balanced across variables of interest.

An average household has about 2 school-age children and is located about 3.1-3.3 kilometres from the child's school. Roughly 38% of households engage in casual labour and about half run household farms.

When it comes to guardian characteristics about 72% are married. The average age is 33 and mothers account for 51% of guardians. Altogether, 92% of guardians are biological parents to the child of interest. **Only 76% of parents are literate** which may impede their ability to participate in the intervention. Anecdotally, we found that Illiterate parents ask their child or another household member to read out Movva SMSs to them[14]. Further, 45% of parents neither make visits to their child's school nor talk to their child's teacher. More than half of parents mention finances as the largest impediment to child school attendance while 27% of parents have no challenges.

Children are on average 10 years old and ~50% in both treatment and control groups are girls. Current school enrolment is very high at 97%, however, **more than half of the selected children (55%) report repeating a grade at least once in their academic career.** We do not have information on the reasons for grade repetition in the baseline data. Conversations with school leaders indicate that grade repetition is in part driven by the parent's inability to pay administrative fees that would allow a child to sit for end of year exams. Children miss school about 0.86 days in a week and spend about 0.9 hours per day helping parents with chores.

---

[14] We examine differences in message comprehension between literate and illiterate participants in table B3, section 3.4.

Table 2: Balance checks on baseline characteristics across treatment and control groups

	(1) Control Mean	(2) Treatment Mean	(3) Difference	(4) N
Household Characteristics				
Distance to school (km)	3.099	3.440	0.341	1839
Household Farms	0.466	0.457	-0.009	2096
Household Does Casual Labor	0.383	0.395	0.012	2096
Number of school-age children in the household	2.211	2.213	0.002	2096
Characterstics of a randomly selected parent				
Age	33.358	34.026	0.668	2096
Married	0.724	0.733	0.009	2096
Parent never visits school	0.455	0.461	0.007	2096
Parent never talks to teacher	0.420	0.418	-0.002	2096
Main challenges with child school attendance				
Sickness	0.049	0.057	0.008	2096
Finances	0.565	0.522	-0.043*	2096
Supplies	0.226	0.200	-0.026	2096
No challenges	0.278	0.276	-0.001	2096
Parent is Literate	0.760	0.770	0.010	2095
Years of education	6.489	6.621	0.132	2095
Respondent is a mother	0.517	0.488	-0.029	2095
Respondent biological parent of child	0.918	0.887	-0.030**	2095
Characterstics of an interviewed child				
Male	0.489	0.500	0.011	2096
Age	10.257	10.324	0.067	2096
Enrolled	0.974	0.979	0.005	2035
Repeated Grade	0.558	0.549	-0.009	2096
Days missed in the past 1 week	0.865	0.866	0.001	1829
Total hours of chores yesterday	0.918	0.858	-0.060	2096

Notes: The child for the interview was randomly selected
Pvalue from f-test on joint orthogonality=0.507
The value displayed for t-tests are the differences in the means across the groups.
Standard deviations are robust.
Fixed effects using variable strata are included in all estimation regressions.
***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

3.2 Takeup and spillovers

The results on the takeup of the messages are presented in **Table 3**. In columns 1-4, we report means in treatment and control group separately, along with the number of observations for each. In column 5, we report the difference between the groups (where

applicable). Statistical significance of the difference is denoted with stars[15]. In Column 6, a total number of observations are reported.

Overall, we see very high receipt of messages in the treatment group, and according to Movva administrative records, 100% of the treatment group was sent messages. The number of messages received varied by wave of the intervention and ranged from 15.8 messages (for the latest wave 3) to 39.8 messages (for the earliest wave 1). **About 90% of the treatment group report receiving at least one message** (Panel B). To our surprise, we also see non-zero reports of receipt of messages in the control group, and roughly 6% of the full control group reports receiving messages directly from Movva. After qualitative follow-ups with these respondents, we do confirm that they were exposed to the intervention, however, it is more likely that they have received messages from others in their community since the SMS implementer did not have access to phone numbers of the control group. Conditional on receiving messages, the treatment group reports receiving 2.32 messages per week, while the control group reports receiving 0.46 less messages. The difference is likely suggestive of the control group only receiving a fraction of the messages that were coming from others in their community.

During data collection, we discovered that Airtel (one of two major telecom companies in Malawi) also sent messages related to education, which some respondents mistook for messages coming from Movva. Airtel messages contained reminders that exam fees could be paid by Airtel money and should not have been a direct substitute to messages sent by Movva. Roughly half of the sample received these messages. The treatment group reports receiving these messages at higher rates, likely due to being more attentive to messages in general, given they are enrolled in the intervention.

In panel C, we examine message spillovers from treatment to control group and observe that there is unprecedented rates of message sharing by the treatment group. We find that **51% of the treatment group report discussing content of the messages with somebody outside of the household.** Conditional on that, the average number of people who the messages are discussed with is 5.2 (distribution of number of people is shown in **Figure 2**). About 18% of treatment individuals report directly forwarding the messages to others, and, if they do so, the number of messages forwarded is about 5.3 (distribution of number of messages sent to others is shown in **Figure 1**). The vast majority of people only forwarded fewer than 4 messages and shared content with fewer than 5 people, though, there is variation.

Given the total number of messages in the intervention is much higher than what is shared with the control group, we don't consider this to be a deep exposure to treatment. However, we cannot rule out that even this level of exposure produced impacts in the control group, thereby reducing the possibility of detecting treatment effects.

Similar patterns of message exposure are observed in the control group. **About 48% of the control group report that the messages are either sent to them directly or discussed by somebody else in the community.** This suggests that the internal validity

---

[15] We use standard notation. *- 10%, ** - 5%, *** -1% significance level.

of the experiment is compromised and the estimate of differences in outcomes of interests between the groups is likely biassed. However, **this level of content sharing likely suggests extremely high enthusiasm about the messages** in communities, which is confirmed by universally positive impressions of the intervention as discussed in Section 3.4. We caveat, however, that this level of enthusiasm may not hold outside this context since the population of interest in the study is unique. First, these communities are rather small and the experiment took place in 32 villages with a total estimated household population of 7,556 and there are likely saving groups and community meetings in which people exchange information with each other. Second, GiveDirectly recipients are likely exchanging information on the benefits they receive with each other more compared to other populations and likely considered it to be part of the same program and cross-validated with each other what other program benefits they were receiving.

Next, we explore the correlates between sharing messages and household and parent characteristics in **Appendix B, Table B1**, and see that compared to guardians who do not share messages, **sharers tend to be female, more likely to be literate and are more educated and more likely to be a biological parent** for child enrolled in the study.

Table 3: Takeup of Movva messages and spillovers across treatment and control groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Control		Treatment			
Panel A. Movva Administrative Data	Mean	N	Mean	N	Difference	N
% of treatment respondents that received at least 1 message			1.0	1176		
Number of messages received (wave 1)			39.8	765		
Number of messages received (wave 2)			22.8	192		
Number of messages received (wave 3)			15.8	219		
Panel B. Survey Data - Message Receipt						
% of respondents that report receiving Movva messages	0.06	1047	0.90	1048	.8396***	2095
Number of messages received per week	1.85	54	2.31	330	.4602***	384
Received messages from Airtel *(asked in Round 2 only)*	0.48	582	0.56	594	.0880***	1176
Panel C. Survey Data - Spillovers						
Treatment (only asked for Round 2 respondents)						
Within HH						
Somebody else in the household checks messages			0.42	594		594
Across HHs						
Somebody checks messages (outside of hh) OR forwards messages to others OR discusses content with others			0.54	594		594
Somebody outside of household checks messages			0.03	594		594
Forwarded messages to others			0.18	594		594
If forwards messages, number of messages forwarded			5.30	108		108
Discussed contents of messages			0.51	594		594
If discusses messages, number of people the messages were discussed with			5.21	304		304
Control *(only asked from a random sample of 250 households in Round 2)*						
Across HHs (only asked for a random sum-sample)						
Someone in community sent messages to them	0.19	220				
Someone in the community discussed messages with them	0.45	220				
Someone in the community either sent or discussed messages with them	0.48	220				

Figure 1: Distribution of the number of
messages forwarded by treatment households
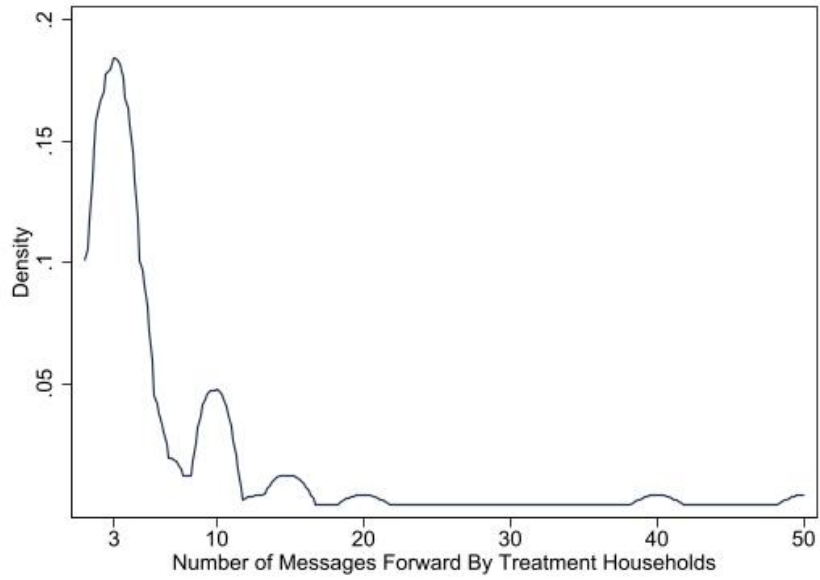
Figure 2: Distribution of the number of people
treatment households discussed messages with
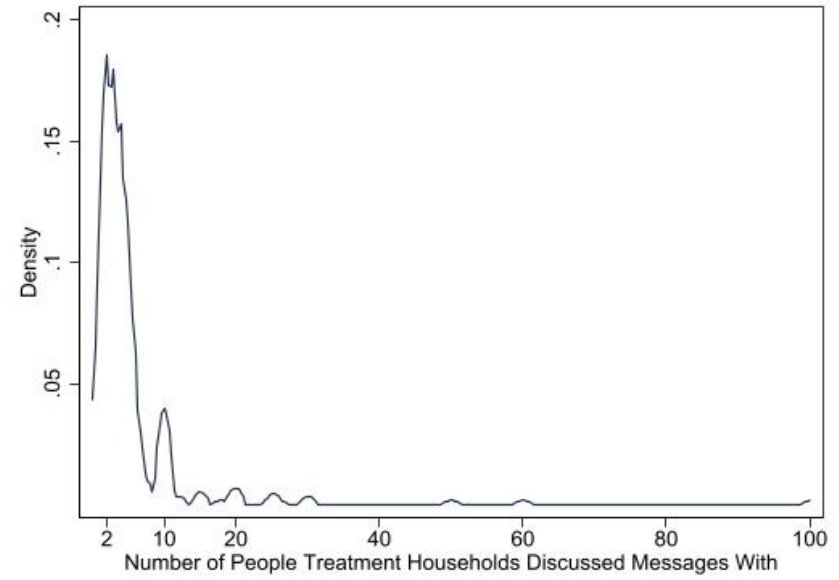
3.3 Causal evidence

3.3.1 Impact on attendance and parental engagement

The results on the impact of the intervention on the pre-specified outcomes are presented in **Table 4**. In columns 1-4, we report the control group mean, the treatment effect, and the p-value associated with the estimate. The control group mean is equal to zero for all indexes as the control group was used as the reference group during the index construction process. In column 5, the total number of observations is reported.

The primary outcome, attendance index, was constructed using child reports of attending school yesterday and the percentage of days (out of 4 data points collected during school visits) child attended school according to school records. We see treatment effects of 0.007 standard deviations, which is very close to zero, suggesting that there are no detectable impacts on attendance. An effect size of this magnitude is very low compared to other studies in education where the medium effect size is 0.1 SD on learning (mean 0.16), to achieve this effect size in learning, one would expect an even larger effect on attendance if it were the main barrier to learning (Evans et al. 2022). We report the treatment effects on self-reported attendance in **Table 5**.

In secondary outcomes, we examine 5 indexes and 2 sub-indexes looking at how a child uses their time, how parents engage with their children, what motivates children when it comes to school, whether children will be enrolled in school and how they performed in the last set of examinations. The time use index was constructed using child reports on the hours they spend doing chores, assisting with farming activities, and with non-farm business activities. We observe a very small point estimate of 0.042 standard deviations, we are unable to reliably conclude that the intervention impacted the hours children spend assisting within the household. To understand how parents and children engage, we constructed the child-parental engagement index with two sub-indexes for general engagement and school related engagement. The general engagement sub-index was constructed using child reports of how often parents engage with their child[16]. Like the general engagement sub-index, the school sub-index was constructed from child reports on how often parent exhibit a certain behaviour[17]. For the overarching index we see treatment effects of 0.047, mainly driven by the general engagement sub-index with treatment effects of 0.052 while the school engagement sub-index has effects of 0.036, and none are statistically significant.

The progression score index was constructed using data collected from progress reports

---

[16] The index was constructed of the following behaviours: how often the parent plays with their child, tells their child stories, does fun recreational activities with their child, asks the child about their feelings or helps their child manage their fears

[17] How often parents ask the child about their homework, grades, school day, and encourage their child to study.

during school visits. We note that we are limited to complete this analysis for the full sample because more than half of the students were not found in school records[18] (a more detailed explanation is provided in Section 3.5.2). Since each teacher records final scores differently and sometimes the same teacher doesn't use a consistent grading rubric[19], we first converted all scores into comparable units within each grade, and then standardised scores of treatment students against the control students in the same grade and school to make sure scores were comparable across grades. If there were not enough children within grade/school to generate standardised scores, those children were dropped from the analysis. Overall, we find no significant impact from the intervention on progression scores. We note that the lack of data in some cases limited our ability to draw firm conclusions.

The child motivation index was constructed using child reports to why the child goes to school and why the child does their homework. As with other outcomes of interest, this one is not statistically significant. In the last row we report impacts on the probability of enrollment in the January 2023 term. Since we had spoken with households in December, we asked whether the child planned on being enrolled. In the same nature as the previous indexes we observe a small treatment effect and a p-value that does not allow us to conclude that the intervention impacted enrollment.

Overall, **we do not have conclusive evidence that the intervention impacted any of the outcomes of interest which is either due to lack of statistical power to detect very small effects or presence of the spillover effects, or both.**

---

[18] On average, there is about 9 children per treatment arm per grade, with some grades having only one child which was dropped in the analysis (unreported in tables).
[19] Teachers used different grading systems: for some they used the GPA system, for others they used the average mark across all course, for others they simply used the total mark for all the courses and in some cases they used different combinations of the three

Table 4: Impact of Movva messages across treatment and control groups

	(1) Mean in Control Group	(2) Treatment	(3) p-value	(4) N Observations
Primary Outcome:				
Attendance Index	0.00	.0070	0.87	2073
Secondary Outcomes (Indexes):				
Child time use	0.00	.0425	0.33	2086
Child-parental engagement (general and school)	0.00	.0478	0.26	2086
General engagement sub-index	0.00	.0524	0.22	2086
School engagement sub-index	0.00	.0362	0.40	2057
Progression Score (according to school data)	0.00	.0142	0.91	665
Children motivation	0.00	-.044	0.32	2059
Children re-enrollement In the new term	0.00	.0116	0.77	2086

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

3.3.2 Impacts on disaggregated behaviours

Now, we turn to components of the constructed indexes to understand whether any behaviours individually were changed by the intervention. The results on distinct parental engagement according to child reports are presented in **Table 5**. We report the mean of the control group, the treatment effect of the intervention, the q-value[20] and the total number of observations in columns 1-4. We report child reports on school outcomes and on parental engagement (general and school engagement). While we do not see impacts on any of the tested outcomes, a few means in the control group are interesting to note.

First, **we see very high levels of self-reported attendance of 90%**, suggesting that there may be little room for the intervention to increase it further. For a random subsample of children who report not attending school, follow-up interviews show that 47% of them abstained due to illness, and 19% due to having an unclean uniform or no stationary (see Table B2). We don't see evidence that absence of a clean uniform is correlated with reporting financial difficulties at baseline (Table B4). In this context, not having the right supplies disqualifies the child from participating in school activities and the child is sent back home. Since this is the second most prevalent reason for skipping school, incorporating messages encouraging parents to ensure their child always has a clean uniform and is equipped with stationary could potentially help increase attendance. The high self-reported attendance rate may have been driven by social-desirability bias, the possibility of which we discuss in Section 3.5.3.2.

Second, we see that control children perform about 0.96 hours of chores per day. We find this result surprisingly low given that children are a pivotal part of the household labour force. This result is inconsistent with what we found in available census data which reported that family responsibilities and employment are among the top reasons for primary school children to drop out of school[21]. A potential explanation for what we are observing could be that the children are too young with an average age of roughly 10 years old so they are not yet at an age to take on more responsibilities around the house.

---

[20] An equivalent of p-value after adjustments for multiple hypothesis correction
[21]http://www.education.gov.mw/index.php/edu-resources/documents-and-publications/category/10-reports?download=28:2021-malawi-education-statistics-report

Table 5: Treatment effects on school outcomes and disaggregated parental engagement behaviours

	(1) Mean in Control Group	(2) Treatment Effect	(3) q-value	(4) N Observations
School Outcomes				
Went to school yesterday	0.90	.0068	0.93	2070
Total hours spent doing homework yesterday	0.76	.0115	0.95	2086
Total hours spent doing chores yesterday	0.96	.0185	0.93	2086
Arrived to school on time	0.97	-.011	0.80	1041
Parental Engagement (1=Never, 5=Always)				
General engagement				
Played with the child	2.59	.0856	0.68	2080
Read the book	2.75	-.006	0.95	2079
Sang a song	2.61	.0068	0.95	2081
Told a story or asked to tell a story	2.58	.0226	0.93	2078
Did fun activites	2.51	-.042	0.93	2076
Asked about feelings	2.88	.0838	0.67	2079
Asked about fears	2.39	.1013	0.67	2080
Help manage fears and difficulties	2.49	.0406	0.93	2076
School engagement				
Frequency of helping with school work	2.83	.0403	0.93	2053
Asked about school work	3.16	.0863	0.67	2053
Asked about grades	3.46	.0215	0.93	2052
Asked how the day went	3.20	.0197	0.93	2052
Encouraged not to be late	4.16	.0027	0.95	2053
Encouraged not to miss class	4.14	.0093	0.95	2052
Encouraged to study or read	3.98	.0046	0.95	2053

Notes: q-values are reported after the application of FDR multiple hypotheses adjustment. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

3.3.3   Impacts on punitive behaviour and aspirations

The results on parental outcomes on views on punishment and aspirations for their child are presented in **Table 6**. We constructed an index consisting of the following behaviours parents engaged in in the past 1 month to discipline children: Privileges being taken away, physical punishment, or use of demeaning words. As with other outcomes, we see no impacts on either views or behaviours towards punishment or aspirations as indicated by high q-values.

It's worth noting that the self-reported prevalence of these behaviours is low (though, we

cannot exclude social-desirability bias).[22] Taking away a child's privileges is the most prevalent form of punishment used by about 23% of households. Prevalence and acceptance of physical punishment is reported to be relatively low at 6% engaging in the behaviour and 12% endorsing the behaviour[23]. Given the sensitivity of the topic, we avoided asking children directly about parental abusive practices, so the parental reports may be lower than the actual occurrence of abuse experienced by a child. For example, the Malawi 2019-20 Multiple Indicator Cluster Survey found that while only 15.5% of caregivers considered physical punishment to be an acceptable method of discipline[24], 62% of children reported being subjected to psychological or physical punishment in the past 1 month, 14.1% experienced a non-violent disciplinary approach, 73.5% experienced psychological aggression, and 17% faced severe physical punishment. This suggests that despite parents having negative views on punishment, children actually experience it at relatively high rates.

Lastly, we note that parents report unusually high educational aspirations for their children as 67% hope that their child will obtain more than 12 years of education.

Table 6: Parental reports on views towards punishment and educational aspirations for their child across treatment and control groups

	(1) Mean in Control Group	(2) Treatment	(3) q-value	(4) N Observations
Views towards punishment				
Anderson index on punishing behaviours	0.00	.0187	0.93	2094
Privelleges were taken away (0=No, 1=Yes)	0.23	-.001	0.95	2094
Physically punished (0=No, 1=Yes)	0.06	.0033	0.93	2094
Shouted by any adult member (0=No, 1=Yes)	0.20	.0142	0.93	2094
Called demeaning words (0=No, 1=Yes)	0.08	.0016	0.95	2094
Physical punishment acceptable (0=Never,	0.12	.0344	0.44	2094
Aspirations				
Real aspirations: higher than upper secondary	0.67	-.015	0.93	2094

Notes: The estimations are [control variables]. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

---

[22] To reduce the possibility of social desirability bias, we asked the parent to report whether the child was punished by any household member (not the interviewed parent), which may have helped induce more honest reporting, however, we recognize that there may be reluctance to underreport this type of behaviour in general.

[23] Our findings on acceptance of physical punishment contrast with other studies done in Malawi examining physical child abuse, and Blaney et al. 2019 find that a significant proportion of caregivers (87.6%) found physical punishment to be an acceptable form of discipline, with 73.6% of them engaging in the behaviour (Blaney et al. 2019). The discrepancies in findings are likely driven by differences in measurement and in Blaney et al. 2019 the parents were presented with vignette scenarios which they rated as acceptable/not acceptable. This measure might have elicited more honest responses from caregivers.

[24] https://washdata.org/sites/default/files/2022-02/Malawi%202019-20%20MICS.pdf

3.3.4   Heterogeneity in treatment effects

For this analysis, we used "tree"-based models, which can be intuitively understood as a kind of matching of households with similar characteristics into buckets where a treatment effect is then estimated for that bucket. This is helpful when we do not wish to specify how different variables might interact together to influence the relationship between the treatment and outcome. Multiple trees can be grouped together to refine such estimates. For a more detailed explanation of the methodology, we follow Athey et al (2019)[25] for algorithm of generalised random forests and Athey et al (2020)[26] and Chernozhukov et al (2017)[27] for methodology of applying causal forests to heterogeneous treatment effects, specifically on quantifying heterogeneity. We use a double machine learning implementation of causal forest by Microsoft EconML[28] for Python. For the cross validation procedure, we follow Chernozhukov et al (2017)[29]. For feature selection, we used baseline characteristics[30] specified in Section 2.4 as controls with additional variables, and tested for potential heterogeneity.

For estimation of heterogeneous treatment effects, we look at two outcomes: child reported attendance and the parental engagement index. The training data consists of 966 randomly selected households with another 415 households withheld as a test set.

While we observe possible variation in treatment effect across certain subgroups (defined by variables above), the differences between the conditional average treatment effects of these groups are not statistically significant.

Although there appears to be some variation in treatment effects across households based on selected variables (particularly when demarcated by child age and distance from school), **we are unable to detect any statistically significant subgroups**. In addition, **we find that household characteristics are not strong predictors of either parental engagement or child reports of attendance, which suggests that included variables for analysis likely do not exert a strong effect on the outcome.** In **Figure E1**, we present the results of a validation exercise to check whether the variation in ATEs are significant. The x-axis depicts quartiles for predicted average treatment effects and the

---

[25] Athey, Susan, Julie Tibshirani, and Stefan Wager. (2019) Generalised random forests. The Annals of Statistics 47.2 (2019): 1148-1178. (https://arxiv.org/pdf/1610.01271.pdf)
[26] Athey, S., Keleher, N., & Spiess, J. (2020). Machine Learning Who to Nudge: Causal vs Predictive Targeting in a Field Experiment on Student Financial Aid Renewal
[27] Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I.. (2017). Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomised Experiments. (https://arxiv.org/abs/1712.04802.pdf)
[28][EconML](https://econml.azurewebsites.net)
[29] The following procedures were used (1) Split the data into K folds (2) for each fold k, train a CATE estimator G on the other K - 1 folds and predict on k. Then create a mapping Q: x → {1,2,3,4} where Q(x) := CATEQuartile(G(x)) i.e. map a point to the quartile its CATE estimate belongs to using the K - 1 folds. Then apply Q to fold k. (3) Finally, for each quartile, estimate the sample ATE and SE
[30] Baseline attendance, Gender of the child, Grade level of the child, Distance to school, Gender of the interviewed parent (targeting female vs male recipients may have different impacts), Literate vs illiterate guardians as SMS may have been less successful for parents who cannot read/write, Whether parent reported to ever visit school at baseline (proxy to interest in school, parents with low interest may be affected more), Number of total hours that child helps with chores / farming activities at baseline, Whether baseline absence was due to financial challenges, illness, or other categories, Additional controls corresponding to phase / wave of intervention

y-axis is the estimated sample treatment effect for that quartile. Since we see that the error bars overlap across each quartile of predicted treatment effects, this suggests that we cannot conclude that variations in treatment effects are not just due to noise.

With respect to model performance, having strong predictors is important in reducing the error in the estimates and improving precision of findings. In ***Figure E2*** we present a decision tree that first splits by number of school aged children in the household, then time taken on chores, to predict an average treatment effect in each bottom subgroup (at the bottom of the figure). Even though we might have identified consistent splits, the confidence intervals driven by model error within each subgroup includes zero. **This means we cannot distinguish the effect within each group from the others/conclude that it is significant.**

3.4 Descriptive evidence

3.4.1 Comprehension of messages among treatment group

In this section we investigate whether the messages were understood by the respondents. The results on parents comprehension of messages are presented in **Table 7**. The table details only responses from treatment households that report receiving messages from Movva. In Panel A, we see that 93% of parents report understanding the content of the messages. In the follow up question, we asked the respondents to highlight their main takeaways of the message content. These free responses given by guardians suggest a high rate of message comprehension: 74% of parents easily recall messages encouraging them to talk to their children about how school is going.

We further asked which messages they find most interesting and parents report finding messages reminding them to encourage their child to go to school (39%), encourage their child's future ambition (20%), and to continue providing necessary educational materials (18%) as the most interesting.

All parents report interest in receiving these messages in the future which is not surprising as roughly 97% of them report the messages to be either very useful or extremely useful (Panel B).

These self reports show a positive reception, interest and excitement for Movva messages, however, it is possible that the short-term excitement may subside in the medium or long term.

One barrier to participation in the intervention may be ability to read. To explore this possibility, we compare experiences with Movva messages of literate and illiterate caregivers in the treatment group. The results are presented in **Table B3**. We find that while some illiterate respondents are able to engage with the messages, **measures of**

engagement are higher for literate respondents. Specifically, literate respondents are more likely to report receiving messages (14% points difference), more likely to understand the content of the messages (20% point difference), more likely to forward messages to other people (7% point difference), more likely to discuss messages with other people (23% points difference). This suggests that inability to read negatively impacts engagement with the content.

Table 7: Parental comprehension of Movva messages for treatment households who report receiving messages

	For treatment group only	
	(1)	(2)
	Mean	N
Panel A: Message Compresion		
For those who report receiving messages		
Respondent was able to understand the content of the messages	0.93	858
Content of the messages		
Encourage parents to talk to children about how school is going	0.74	858
Emphasize the importance of education	0.37	858
Provide tips on how to build good relationships with their children	0.22	858
Encourage parents to help children with their homework	0.22	858
Advise parents on the best ways to discipline children	0.19	858
Encourage parents to talk to children about non-education topics	0.11	858
Don't remember	0.02	858
Messages that they found most interesting/useful		
A message remindng parents to encourage child(ren) to go to school	0.39	854
A message to parents to encourage child(ren)'s future ambitions	0.20	854
A message for parents to continue providing necessary education materials for school	0.18	854
A message about parent's awareness of their child(ren)'s performance in school	0.17	854
A message reminding parents to know their child(ren)'s teachers(s)	0.11	854
A message advising parents not to physically punish/yell at child(ren)	0.08	854
Don't remember	0.11	854
Panel B: Future receipt		
Interested to receive the messages in the future	1.00	858
Messages are useful		
Somewhat useful	0.01	858
Moderately useful	0.02	858
Very useful	0.48	858
Extremely useful	0.49	858

3.4.2 Self-reported impacts of messages

To triangulate experimental findings, we administered a short section on self-reported impacts of the intervention, the results of which are presented in **Table 8**. The table describes if and how parents changed their behaviour towards their child in Panel A and if and why their child's attendance changed in Panel B.

The vast majority of parents report being positively affected by messages and **84% of parents report engaging with their child differently after receiving the messages from Movva**. While we cannot fully rule out that these high reports are driven by experimenter demand bias, we have formulated the question to mitigate this concern and allow for

negative reporting[31]. On school related matters, we see that half of parents report now reminding their child the importance of school, 47% now asking their child about homework and school work, and 42% now assisting their child with their homework and school work. Though the prevalence is low (at only 7%), parents report that they now ensure that their child does not go to school on an empty stomach. The latter was not an explicit encouragement of the messages, but indicates that some parents are now more attentive to their children's needs. Parents mention non-school related changes at lower rates, 11% of parents report that they absolve children from chore responsibilities and 8% no longer verbally punish their child. For parents who report not changing their behaviour, we find that 87% of them already had a good relationship with their child before the intervention.

Concerning child attendance, **93% of parents report that their child now attends school more as a result of Movva messages**. 64% of parents attribute this increase in attendance to the child understanding the importance of becoming educated, 42% attribute it to them personally making sure that the child has gone to school, and 42% specify that it is due to discussions about the importance of school. For the 7% of parents who reported that their child did not attend school more, 88% of them have children who were not skipping school before Movva messages began.

---

[31] Example of question asked: There are no right answers here, we want to hear your unique experiences being a parent. Sometimes these types of SMSs change how parents interact with their children and sometimes not. Do you think anything has changed in how you interact with ${child_name}?

Table 8: Self-reported impact of messages - parental behavioural change towards child and child changes in attendance

	For treatment group only	
	(1)	(2)
	Mean	N
Panel A: Self-reported Parental Behavior change		
Parent changed behaviour after receiving messages	0.84	525
What are the ways in which parent changed?		
School		
Remind child importance of school	0.49	374
Ask child about homework and school work	0.47	374
Assist child with homework and school work	0.42	374
Have a change in attitude regarding the importance of school for child	0.30	374
Ask about situations child meets at school	0.22	374
Ask why child doesn't want to go to school and encourge them depending on their reason	0.16	374
Visit child's school	0.10	374
Ensure child doesn't go to school on an empty schomach	0.07	374
Got child tutor	0.06	374
Non-school		
Do no absent child to help with chores	0.11	374
Does not verbally punish child	0.08	374
Ask about how child interacts with friends	0.07	374
Excuse child from doing chores so they can do school work	0.07	374
Excuse child from doing piece work so they can go to school	0.05	374
Why not change behaviour:		
Relationship was already good before messages	0.87	68
Panel B: Self-reported changes in child attendance		
Child attended more school	0.93	526
For those who attend more, reasons why child does attend school more:		
Child understands the importance of becoming educated	0.64	413
Parent ensures child attends school	0.42	413
Parent explained to the child the importance of getting an education	0.42	413
Paren talks to child on why not to miss school	0.28	413
Child feels motivated by the themes addressed by the Movva messages	0.23	413
Due to increased participation by parent in child's education, child has more interest in school	0.16	413
Child more focused	0.12	413
Parent ensures child is ready for school before going to work	0.11	413
Child feels encouraged by examples of role models in the village	0.08	413
Child simply taken more interest in school	0.06	413
Increased availabily of food and clothes	0.05	413
For those who do not attend more, reasons why child does not attend school more:		
Child never skipped school before	0.88	25
Child constantly sick	0.04	25
Child dropped out	0.04	25

3.4.3 Barriers to engagement

Sometimes parents receive information but may be unable to act on it for various reasons, which could help explain the lack of impact we observe in the experimental results. **Table 9** presents the findings on the potential barriers for parents to engage with their child. In the table, we investigate 3 categories of barriers to engagement that could potentially affect the impactfulness of messages, namely: parents may not be able to remember to engage (Panel A) or find time to talk (Panel B) to their children, and, in some cases, children may not positively respond to parental nudges and ignore the messaging (Panel C). **We do not find any of these barriers to be a compelling explanation for unimpactful results**. We find that 78% of parents report always remembering to engage with their child and 81% expressing preference for continuing to receive messages. We also see that 78% of parents always find time to talk to their child; for those that are unable to find time it is mostly due to being busy with work or being too tired. It is encouraging to see that 95% of parents feel that their child's interest in school has increased as a result of the parent's interest.

Table 9: Self-reported barriers for parents to engage with their child

	For treatment group only	
	(1)	(2)
	Mean	**N**
Panel A: Self-reported ability to remember to engage		
Parent rememebered to engage with child		
Always	0.78	526
What can help with remembering better:		
Continue sending messages	0.81	442
Increase message frequency	0.24	442
Send weekly reminders	0.12	442
Include in-person talks	0.11	442
Encourage parent to sit and explain content to child	0.11	442
Incorporate calls with households	0.09	442
Nothing	0.03	442
Change message timing to evening	0.02	442
Panel B: Self-reported ability to find time to talk		
Found time to talk to the child		
Always	0.78	526
Prevents parent from talking to child:		
Busy with work	0.59	98
Too tired after work	0.40	98
Nothing prevents me from talking to my child, I already talk to my child.	0.06	98
Nothing prevents me from talking to my child, I just don't do it.	0.03	98
Child does not live in the household	0.00	98
Panel C: Self-reported changes in child's school interest		
Children's interest in school as a result of parental interest		
Child became more interested in school	0.95	525

3.4.4   Unintended consequence

In some contexts, even small interventions may have unintended consequences that are not anticipated by researchers and small qualitative follow-ups to document those can shed light on areas that are not covered by the research themes. For example, there could be an increase in intra-household disagreements as mothers start to engage with children differently. For this intervention, in **Table 10**, we pleasantly observe a lot of positive self-reported unintended consequences. It is interesting to note that the **themes**

touched on by the messages on how parents should engage with their children seem to also be applied to other adult members in the households. Following the intervention, 19% of parents report that they now talk to other household members about their goals and ambitions, another 18% report they now encourage other household members to pursue their dreams, and 18% report they now make time to listen and engage with other household members.

Table 10: Self-reported impact of messages changes in interaction with other household members

	For treatment group only	
	(1)	(2)
	Mean	N
Changes in interactions with other household members		
Parent encourages them to pursue their dreams	0.19	395
Parent talks to them about their goals and ambitions	0.18	395
Parent makes time to listen and engage wiith them	0.18	395
Parent helps other HH member think through how they can achieve their goals	0.12	395
Parent asks them how their day was and talk about any challenges they may have met	0.10	395
For the younger ones, parent encourage them to go back to school and get an education	0.08	395
Do other things with household members	0.08	395
Parent is more more understanding when they make mistakes	0.05	395
Parent celebrates their successes with them	0.02	395

3.5 Threats to validity

3.6.1 Phone survey completion rates

We present phone survey completion rates and school availability records in **Table 11**. For the phone survey, we were able to interview 90% of our sample of 2,333 with a similar proportion of control group (90%) and treatment group (89%). We were unable to reach 7% of households after following our protocol of attempting to call the household a total of 7 times over 4 days. **Overall, the response rates to phone surveys is high and no-different in both groups.** We do not see differences in reach rates across the randomisation waves; we see a reach rate of 90% for wave 1 and 3, and a reach rate of 91% for wave 2.

3.6.2 Missing data from schools

The rates of finding children records in schools are much lower and we were only able to find 48% of interviewed children in either the class registers or progress report books, suggesting that roughly half of the children could not be located at schools that either

they or their parents reported going to. This result is disappointing and we have conducted interviews with master teachers and assessed the state of records ourselves to form an understanding of why such a large fraction of children was not found.

The reasons for not finding children could be as follows:

First, **some children use different names** which makes it challenging to locate them in school records. For a sub-sample of respondents, we were able to collect updated information prior to school visits **(see Table 11)**, which increased our chances of finding children. In this group, we found 51% of children, compared to 37% of children for the groups that were interviewed during or after school visits. During interviews we asked children which school they went to, the grade they were in and the name they registered at the school which increased our chances of finding the child in the school records.

Second, roughly 20% of children were not found because **attendance registers for the classes of interest were completely missing**, for another 12% of children, the **register was available but not updated** with attendance information for the days of interest.

When it came to test scores, we were able to find scores for 36% of children interviewed. Similar to the attendance registers, the primary reason for not finding children is unknown with only 20% attributable to missing progress report books.

Third, conversations with head teachers during the school visits brought to light a number of reasons that could help explain missing records further. On paper, schools report having systems in place to keep teachers accountable to take attendance, however, in practice, they are not strictly enforced which result in **teachers taking attendance sporadically**, **oftentimes when they believe that the administrator will request to check the registers**. It is expected that school administrators check the registers on a weekly basis, however, conversations with head teachers reveal that they only carry out this task once a month, suggesting that the teacher incentives for accurate record keeping may be low.

Fourth, we learnt that i**t was common practice for teachers to take both registers and progress report books home which sometimes resulted in records being lost**. According to Ministry of Education guidelines, registers and progress report books are to be handed over to headteachers and stored in their offices, however, in a lot of schools headteachers do not have access to offices which makes keeping the records centrally difficult.

Fifth, head teachers explained that though all children who go to a school are supposed to have a record of registration, this is not always the case and some kids attend school without being formally enrolled (though, the prevalence is estimated to be lower than 5%).

Lastly, we examined the differences between children who were and were not found in schools (out of all children who reported to be enrolled both at baseline and endline). The results are shown in **Table B5** and we find differences in characteristics between those two groups. While we do not have a good explanation of what is driving those

differences, the contrast between the groups suggests that the reasons for not finding the children are not random and there are systematic reasons (other than poor record keeping) why these children are not found in school records.Children we found in administrative records were from more literate households with a biological parent who had higher education compared to those we couldn't find. Further investigation would be needed to identify possible explanations for the missing data.

Table 11: Survey attrition, availability of registers and progress reports

	(1)	(2)	(3)	(4)	(5)
	Overall	Control	Treatment	Difference	N
Phoney Surveys					
Percentage of respondents surveyed (out of full sample)	0.90	0.90	0.89	-.012	2333
Percentage of respondents surveyed refused	0.00	0.00	0.00	-.000	2333
Percentage of respondents surveyed not reached	0.07	0.07	0.07	.0023	2333
Percentage of respondents surveyed ineligible	0.03	0.03	0.04	.0123*	2333
Percentage of respondents surveyed (wave1)	0.90	0.90	0.89	-0.01	1513
Percentage of respondents surveyed (wave2)	0.91	0.94	0.88	-0.06	382
Percentage of respondents surveyed (wave3)	0.90	0.90	0.90	0.00	438
In-person data collection (only for children who completed the survey)					
Found a child in school (out of the full sample) either present in register or progress report	0.48	0.48	0.48	-.003	2086
Interviewed before school visits and either present in register or progress report	0.51	0.52	0.51	-.003	1587
Interviewed after school visits and either present in register or progress report	0.37	0.38	0.37	-.010	499
Attendance Records:					
Access to January attendance records (=1 we have access, 0 = no attendance records)	0.43			-.002	2086
For those without attendance records:					
% of kids for who information on school was not provided at BL and EL	0.00	0.00	0.00	-.001	1335
% of kids who were affected by missing attendance registers for relevant classes	0.19	0.20	0.19	-.013	1335
% of kids who were found in registers but attendance was missing	0.12	0.13	0.10	-.028	1335
% of kids who weren't found in schools for unknown reason	0.73	0.71	0.75	.0305	1335
Test Scores:					
% of Kids for which we found a December test score in schools	0.36	0.36	0.36	-.003	2086
For those without test score record:					
% of kids for who information on school was not provided at BL and EL	0.002	0.00	0.00	-.001	1307
% of kids who were affected by missing grade books for relevant classes	0.197	0.20	0.20	-.001	1307
% of kids who were found in gradebooks but attendance was missing	0.015	0.01	0.02	.0176***	1327
% of kids who weren't found in schools for unknown reason	0.801	0.80	0.80	.0027	1307

3.6.3 Self-reports given by children

3.6.3.1 Parent-child validation

Interviewing young children may introduce measurement errors which could reduce reliability of the results. To explore this possibility, we asked children and adults similar questions and present results in **Table 12**. Overall impact estimates and conclusions derived from parental and child reports look very similar. Both parents and children report similar attendance around 90%. However, **engagement reports between children and adults differ and parents tend to be more pessimistic** about their engagement with children compared to children. This might be due to children defining what the engagement looks like differently from how parents would explain it.

Table 12: Parent-child report validation on school outcomes and parental engagement

	Child Reports			Parent Reports		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean in Control Group	Treatment Effect	q-value	Mean in Control Group	Treatment Effect	q-value
School Outcomes						
Went to school yesterday	0.90	.0068	0.91	0.91	.0046	0.92
Total hours of HH chores	0.95	.0241	0.91	0.90	.0242	0.91
Total hours of farming activities	0.34	.0163	0.91	0.24	.0160	0.91
Total hours of non-farming business activities	0.12	.0174	0.91	0.10	.0060	0.96
Parental Engagement in the past 1 week (0=Never, 1=Otherwise)						
General Engagement						
Played with the child	0.69	.0489	0.61	0.65	.0370	0.61
Read the book	0.73	.0137	0.91	0.57	.0355	0.61
Sang a song	0.75	.0340	0.61	0.57	-.006	0.92
Told a story or asked to tell a story	0.75	.0304	0.61	0.55	.0222	0.86
Did fun activites	0.73	-.004	0.96	0.51	.0174	0.91
Asked about feelings	0.84	.0094	0.91	0.66	.0104	0.91
Asked about fears	0.72	.0249	0.69	0.48	.0116	0.91
Help manage fears and difficulties	0.74	.0273	0.61	0.55	.0238	0.82
School engagement						
Helped with school work	0.76	.0059	0.92	0.44	.0335	0.61
Asked about school work	0.88	.0311	0.61	0.64	.0265	0.69
Asked about grades	0.93	.0126	0.79	0.75	.0207	0.82
Asked how the day went	0.87	.0214	0.61	0.58	.0337	0.61
Encouraged not to be late	0.96	.0040	0.91	0.91	.0221	0.61
Encouraged not to miss class	0.96	.0024	0.92	0.91	-.018	0.61
Encouraged to study or read	0.97	-.002	0.91	0.88	.0082	0.91

Notes: The estimations are [control variables]. ***, **, and * indicate significance at the 1, 5, and

3.6.3.2 School records-child report validation

In order to examine validity of children's report on attendance, we turn to administrative data obtained from schools and zoom in on a subsample of 59 children for whom we have overlapping dates for attendance between phone surveys and school records (by

chance those children happened to be asked about days for which we check the registers). Before discussing the results, we note that we don't have full confidence that the school records are completely accurate given our team's observations in the field as well as conversations with head teachers. For example, in examination of registers, we found that it was possible for one child's attendance to be updated (with a present or absent mark), while another child's in the same register attendance was simply left blank for the exact same day. **Conversations with head masters confirmed that teachers are not consistent in updating registers daily**, and instead, teachers may fill in attendance for all days at once right before the registers are checked by the school administrator (once a month) and Primary Education Advisor (2-3 times a term), relying on memory and recall. This suggests that discrepancies between children self-reports and attendance records may not necessarily all be due to children misreporting but could be due to inaccurate school records.

Nevertheless, we present the match rate in school records and self reports in **Table 13**. We find that reports match for the majority of the children (76%); for 17% of children, reports of attendance are inflated; while for 7%, the reports are deflated. In the absence of systematic misreporting the rates of under and over reporting should have been the same. These reporting discrepancies are similar for both treatment and control groups although there is some indication that treatment reports were a little more inflated (18% of reports in treatment group and15% in the control group, respectively). However, we cannot conclude with certainty that these are statistically significant from each other as the sample size is too small. **This suggests that there is indicative evidence that children are systematically over-reporting attendance, but not sufficient evidence to conclude whether this behaviour is different for the treatment group.**

Next, we examine the rates of attendance separately for each day in January. For January 23 2023 which is a Monday, we find a very big difference in attendance rates as 90% of children report attending school while school records show that only 73% of them were actually in attendance. The dissimilarity is alarming and brings to question the accuracy of the child self-reports but also the quality of record keeping in schools as a 17% difference is not easy to explain. Interestingly, the differences between reporting rate is almost negligible for January 25th, 2023 (Wednesday) where 93% of children report attending, and 90% are marked as present on that day. This discrepancy between Monday and Wednesday may be further evidence that **attendance records may not be filled out daily** and accuracy fluctuates depending on whether the teacher actually took attendance or not. Given such a high match rate for Wednesday and low match rate for Monday, there is a possibility that the teachers actually took attendance on Wednesday, and filled in attendance rates based on memory for Monday. However, it is difficult to reach definitive conclusions with such a small sample.

Table 13: Child report and administrative validation of attendance results

	(1)	(2)	(3)	(4)	(5)
	Overall	Control	Treatment	Difference	N
% of kids who's responses match between school attedance data and self reports	0.76	0.77	0.76	-.011	59
% of kids who report attending, but school records specify that they were absent on that day	0.17	0.15	0.18	.0279	59
% of kids who report they didn't attend but school records specify they were present on that day	0.07	0.08	0.06	-.016	59
Monday January 23rd 2023:					
Administrative attendance rate	0.73	0.75	0.71	-.035	30
Self-reported attendance rate	0.90	0.94	0.86	-.080	30
Wednesday January 25th 2023:					
Administrative attendance rate	0.90	1.00	0.84	-.157	29
Self-reported attendance rate	0.93	0.90	0.95	.0473	29

Notes: The number are reported only for the sub-sample of kids where there were overlapping days between phone data collection and dates for which attendances was collected in schools.

3.6.4 Location of study participants

The map of the location of the household is shown in **Figure 3.** The quantification of household proximity is presented in **Table 14**. Panel A shows the number of households within different distances from each other, regardless of treatment assignment. Panel B shows the number of control households within different distances of treatment households. Panel C shows the number of treatment households within different distances of control households. It is important to note that these distances are euclidean distances between geographic coordinates and do not take into account terrain.

Our findings reveal that treatment and control households live very close to each other. On average, each household has 96.6 households within a 1km radius from it. For an average control household, we see that there are roughly 49 treatment households within a 1km radius of the household. We observe a similar situation for the number of control households within a 1km and 2km radius of treatment households. These results are also displayed in **Figure 3**.

The high number of different treatment arm households in close proximity to each other is concerning and helps explain the observed spillover effects discussed in Section 3.2.

Table 14: Number of households within varying distance overall and by treatment arm

	(1) Overall	(2) Control	(3) Treatment	(4) N
Panel A: Number of households within:				
1km	96.6			2096
2km	241.9			2096
Panel B: Number of control households within (x) of treatment household:				
1km			47.9	1049
2km			120.0	1049
Panel C: Number of treatment households within (x) of control household:				
1km		48.6		1047
2km		121.6		1047

Figure 3: Location of study participants

GiveDirectly Movva Education Impact Evaluation
Sample Distribution

Legend
- Treatment Households
- Control Households

0   5   10 km
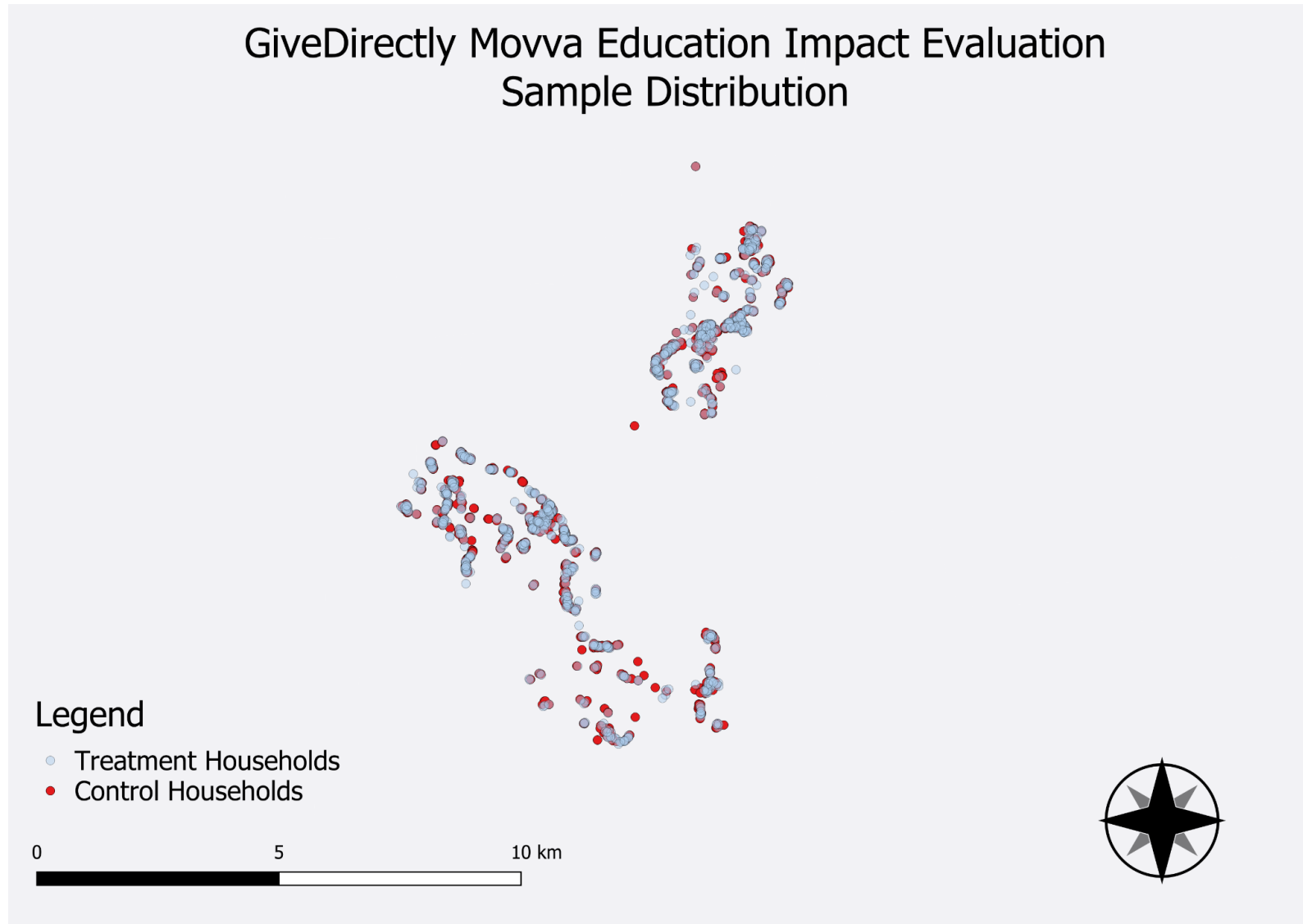
3.6.5   Spillover Effects

To quantify the extent of the spillover effects from treatment to control group, we use variation in saturation of treatment households within a certain radius of control households. Conditional on population density, the saturation rate should be exogenous to outcomes as it's generated by the random assignment which is agonist of the household location. The treatment saturation variable is constructed for each household for a radius of 1,2 and 3 kilometres and represents a fraction of treatment households compared to the total number of households in the area. Following the Miguel et al. 2004, we implement the following equation for estimation of spillover effects

$$Y_i = \beta_o + \beta_1 T_i + \beta_4 PT_i + \beta_5 PT_i * T_i + \beta_6 NHH_i + \beta_2 Y_o + \beta_3 X_i + \delta_1 M_i + \alpha_i + \gamma_t + \varepsilon_i$$

Where:

Y_i denotes the outcome variable for child i

T_i denotes the treatment status of household i (1 for Treatment Group; 0 for Control Group)

NHH_i - number of total households within X km

PT_i - percentage of treatment households (as a function of total households in the sample) within X km. The variable ranges between 0 and 100.

Y_o denotes the outcome variable for child i at baseline. X_i baseline characteristics - These are included to improve precision of the impact estimate.

When X_i is missing for an observation, we include an indicator term for missingness M_{io} and replace X_i with its 0

α_i strata fixed effects

γ_t week fixed effects

ε_i error term

In the above equation, β_1 - effect of the treatment on engagement index given 0 % of treatment HHs with the area ; β_4 - effect of treatment saturation on engagement in the control group; β_5 - additional treatment effect on outcome for every percentage point increase in saturation of treatment households; β_6 - effect of the number of total people within X km away on the outcome of interest.

We report the results in appendix **Tables B6** for 1, 3 and 5 kilometre radius, for 2 variables of engagement (school and general engagement in columns 1-6) and children reports of attendance (columns 7-9). In columns 1, 4 and 7 we report base regressions and gradually add control variables to example how the coefficient changes and whether additional variables have

statistically significant impact on the outcome. We see *indicative* evidence that higher saturation of treatment households increased attendance for 3 kilometre radius specification, and for every percentage point increase in saturation, attendance increases by 0.01 (or 1 percentage point). This is a large effect size and suggests that for every 10% point increase in treatment saturation, attendance increases by 10% points. **This is consistent with the qualitative reports on spillovers and suggests that contamination of the control group is very likely driving the null effects**. We also caution that given the number of hypotheses that are tested in the table, it's possible that this one is significant by chance.

4. Suggested Next Steps if GiveDirectly Chooses to Scale-up

4.1 Adjustments to the intervention

- We recommend **updating SMS content to better meet community needs within their context**. For instance, Movva could include messaging around the importance of having clean uniforms as well as information on how to prevent most common child illnesses in that area to address reported primary reasons for non-attendance in this context– having dirty uniforms and being sick. The latter would require an understanding of prevalent illnesses in this area.

- **Consider conducting qualitative interviews with a subset of the target population to gain a contextual understanding of their challenges in order to better tailor the intervention to their specific needs,** since Movva messages didn't address the primary reasons for non-attendance in this context. The same interviews can be conducted on a yearly basis and contents of the messages can be tailored to the unique contextual factors.

- Movva could consider **further optimising message delivery times for high message access rates**. With 24% of guardians being illiterate, the time of delivery could be shifted to ensure illiterate guardians have a reader close by at time of receipt. This time would still have to be balanced against when most parents tend to be available to read texts.

4.2 Continued implementation monitoring and evaluation for relevance

- Build an M&E plan, select important indicators which define success of the implementation, pre-specify levels for the success and use periodic surveys to track whether to continue or stop the intervention:
    - **Develop a detailed theory of change which highlights important assumptions on the path to impact**
    - **Identify key indicators along the theory of change for measurement.** Aside from attendance, Givedirectly could pre-specify other indicators and levels which determine whether the intervention is working (e.g. 60% of parents are able to recall the content of the last message, zero percent of children report that the reason for skipping school is unclean uniforms etc).
    - **Conduct period surveys with a representative sample of the recipients to track implementation and interest in the intervention.** GiveDirectly could conduct periodic follow-up surveys with recipients to assess the need for messages. While we saw tremendous interest in the messages during this evaluation, it is unclear whether it will be sustained in the medium term. Once interest in the message assuages, the intervention can be discontinued. The intervention can be discontinued if the key indicators do not meet pre-specified targets.

5.   Appendices

Appendix A: References

Angrist, N., Peter, B., & Moitshepi M., (2021). *School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic*. National Bureau of Economic Research. Working Paper No. w28205.

Beaman, L., Ariel, B., Jeremy, M., & Ahmed M.M., (2021). *Can network theory-based targeting increase technology adoption?* American Economic Review, 111(6), 1918-1943.

Benjamini, Y., Abba M.K., and Daniel Y., (2006). *Adaptive linear step-up procedures that control the false discovery rate*. Biometrika, 93(3), 491-507.

Bergman, P., (2019). *How behavioral science can empower parents to improve children's educational outcomes*. Behavioral Science & Policy, 5(1), 52-67.

Blaney, E., Zulu, M., Yiwombe, I., Lipunga, G., Miller, A., Klabbers, G., Feron, F. and Kennedy, N., (2019). *G268 Knowledge, attitude and practice towards physical child abuse in malawi: an explorative quantitative study*. International Child Health Group. doi: 10.1136/archdischild-2019-rcpch.260.

Bureau of Economic Research., (2020).

Cunha, N., Guilherme, L., Ricardo, M., & Eric, B., (2017). *What is it about communicating with parents?* [Unpublished manuscript]. Stanford University.

Doss, C., Erin M.F., Susanna, L., & Benjamin, N.Y., (2019). *More than just a nudge: Supporting kindergarten parents with differentiated and personalized text messages*. Journal of Human Resources, 54(3), 567-603.

Gerber, S.A., & Green, P.D., (2019). *Tailoring instruction to students' learning levels to increase learning*. Journal of Development Economics, 99(2), 210-221. https://doi.org/10.31485/pi.2522.2019

Gertler, P., Heckman, J., Pinto, R., Chang, S., Grantham-McGregor, S., & Vermeersch, C., (2021). *Effect of the Jamaica Early Childhood Stimulation Intervention on Labor Market Outcomes at Age 31*. National Bureau of Economic Research. Working Paper No. w29292. doi: 10.3386/w29292

Guest, G., Namey, E., & Chen, M., (2020). *A simple method to assess and report thematic saturation in qualitative research*. PLOS ONE, 15(5), e0232076. doi: 10.1371/journal.pone.0232076

Lichand, G. and Wolf, S., (2022). *Measuring Child Labor: Whom Should Be Asked, and Why It Matters*. doi: 10.21203/rs.3.rs-1474562/v1

Munje, P. N., & Mncube, V., (2018). *The development and implementation of a mathematics teacher professional development programme in South Africa*. Perspectives in Education, 36(1), 80-93. DOI: http://dx.doi.org/10.18820/2519593X/pie.v36i1.6

Malawi National Statistical Office, (2021). *Multiple Indicator Cluster Survey (MICS) 2019-20 - Survey Findings Report*.

Wager, S., & Susan A., (2018). *Estimation and inference of heterogeneous treatment effects using random forests*. Journal of the American Statistical Association, 113(523), 1228-1242.

Walker, S.P., Susan, M.C., Amika, S.W., Rodrigo, P., James, J.H., & Sally, M. G., (2021). *Cognitive, psychosocial, and behaviour gains at age 31 years from the Jamaica Early Childhood Stimulation Trial*. Journal of Child Psychology and Psychiatry. doi: 10.1111/jcpp.13499.

York, B.N., Susanna, L., & Christopher, D., (2019). *One step at a time: The effects of an early literacy text-messaging program for parents of preschoolers*. Journal of Human Resources, 54(3), 537-566.

Appendix B: Tables

Table B1: Characteristics of treatment households sharing and discussing messages with others

	(1) Mean for Parents Who Did Not Share Content	(2) Difference	(4) N Observations
Household Characteristics			
Household Farms	0.44	.0094	521
Number of school-age children in the household	2.27	-.072	521
Distance to school (km)	3.23	.3528	457
Household Does Casual Labor	0.39	.0008	521
Average distance to control household	8.91	.1965	492
Characterstics of a randomly selected parent			
Female	0.59	-.078*	521
Age	34.14	-2.14	521
Parent is Literate	0.75	.0859**	521
Years of education	6.37	.9149***	521
Parent never visits school	0.49	-.031	521
Parent never talks to teacher	0.43	-.002	521
Main challenges with child school attendance			
Sickness	0.05	-.003	521
Finances	0.52	.003	521
Supplies	0.19	.0204	521
No challenges	0.28	-.019	521
Respondent biological parent of child	0.86	.0610**	521
Aspirations of randomly selected parent for child			
Parent would like child to reach Primary School	0.01	-.009*	521
Parent would like child to reach Lower Secondary	0.00	-.001	521
Parent would like child to reach Upper Secondary	0.13	-.022	521
Parent would like child to reach Post Secondary	0.12	.0200	521
Parent would like child to reach Technical College	0.13	-.007	521
Parent would like child to reach Tertiary Education	0.60	.0217	521
Parent realistically believes child will reach Primary School	0.05	-.022	521
Parent realistically believes child will reach Lower Secondar	0.05	-.000	521
Parent realistically believes child will reach Upper Secondar	0.20	.0432	521
Parent realistically believes child will reach Post Secondary	0.22	-.006	521
Parent realistically believes child will reach Technical Colle	0.11	-.009	521
Parent realistically believes child will reach Tertiary Educat	0.36	-.004	521

Notes: The estimations are [control variables]. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table B2: Reasons for child not attending school yesterday

	(1) Mean	(2) N
Reasons for not attending school yesterday - only asked of a subsample		
Child was sick	0.47	85
No clean uniform or stationary	0.19	85
Funeral in community	0.09	85
Looking after/visiting relatives	0.07	85
Heavy rains	0.07	85
Parent's divorce	0.01	85
Dropped out	0.02	85
No reason	0.02	85
Running other errands	0.01	85
Tired (school far)	0.01	85
Chores	0.01	85
School holiday	0.01	85

Table B3: Ability to engage with Movva messages by literacy (treatment only)

	(1) Mean for Parents Who Are Illiterate	(2) Mean for Parents Who Are Lliterate	(3) Difference	(4) N Observations
% of respondents that report receiving Movva messages	0.797	0.937	0.140***	1048
% of respondents that report other household member receiving Movva messages	0.254	0.270	0.016	594
Received messages from Airtel *(asked in Round 2 only)*	0.575	0.561	-0.014	594
Forwarded messages to others	0.127	0.200	0.073**	594
Discussed contents of messages	0.328	0.565	0.237***	594
Respondent was able to understand the content of the messages	0.768	0.970	0.202***	858
Parent changed behaviour after receiving messages	0.788	0.852	0.064	525
Child attended more school	0.929	0.934	0.005	526
Parent always remembers to engage with child	0.758	0.780	0.022	526
Parent always has time to talk to child	0.778	0.775	-0.003	526
Messages are somewhat useful	0.013	0.010	-0.003	858
Messages are moderately useful	0.040	0.018	-0.021	858
Messages are very/extremely useful	0.947	0.972	0.025	858

Notes: The child for the interview was randomly selected

Pvalue from f-test on joint orthogonality=00

The value displayed for t-tests are the differences in the means across the groups.

Standard deviations are robust.

Fixed effects using variable strata are included in all estimation regressions.

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

Table B4: Reasons child missed school for households that report financial challenges being a barrier at baseline

	(1) Parent Reports Not Having Financial Challenges Mean	(2) Parent Reports Having Financial Challenges Mean	(3) Difference	(4) N
Reasons for not attending school yesterday - only asked of a subsample				
Child was sick	0.558	0.372	-0.186	86
No clean uniform or stationary	0.186	0.209	0.023	86

Notes: The child for the interview was randomly selected

Pvalue from f-test on joint orthogonality=00

The value displayed for t-tests are the differences in the means across the groups.

Standard deviations are robust.

Fixed effects using variable strata are included in all estimation regressions.

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

Table B5: Comparison of household and child characteristics for those found in administrative records (enrolled only)

	(1) Child Not Found In Records Mean	(2) Child Found In Records Mean	(3) Difference	(4) N
Household Characteristics				
Distance to school (km)	3.361	3.068	-0.293	1787
Household Farms	0.444	0.507	0.062***	1979
Household Does Casual Labor	0.413	0.356	-0.056**	1979
Number of school-age children in the household	2.215	2.197	-0.018	1979
Characterstics of a randomly selected parent				
Age	35.102	33.758	-1.344*	1979
Married	0.692	0.764	0.073***	1979
Parent never visits school	0.460	0.446	-0.014	1979
Parent never talks to teacher	0.422	0.406	-0.016	1979
Main challenges with child school attendance				
Sickness	0.049	0.057	0.009	1979
Finances	0.502	0.585	0.083***	1979
Supplies	0.210	0.220	0.010	1979
No challenges	0.283	0.277	-0.006	1979
Parent is Literate	0.734	0.807	0.072***	1979
Years of education	6.333	6.865	0.532***	1979
Respondent is a mother	0.488	0.520	0.032	1979
Respondent biological parent of child	0.889	0.921	0.032**	1979
Characterstics of an interviewed child				
Male	0.495	0.517	0.022	1979
Age	10.630	10.517	-0.113	1979
Enrolled	1.000	1.000	N/A	1979
Repeated Grade	0.590	0.559	-0.031*	1979
Days attended in the past 1 week	0.882	0.895	0.013	1774
Total hours of chores yesterday	0.939	0.825	-0.114**	1979

Notes: The child for the interview was randomly selected
Pvalue from f-test on joint orthogonality=00
The value displayed for t-tests are the differences in the means across the groups.
Standard deviations are robust.
Fixed effects using variable strata are included in all estimation regressions.
***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table B6: Quantification of spillover effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Parent General Engagement			Parent School Engagement			Child Attended Yesterday		
Panel A: Radius of 1km									
Treatment	0.04	0.04	-0.55	0.02	0.02	-0.18	0.01	0.01	-0.06
	(0.36)	(0.35)	(0.13)	(0.66)	(0.71)	(0.65)	(0.70)	(0.69)	(0.62)
% of treatment households within 1 km		0.00	0.00		-0.01	-0.01		0.00	0.00
		(0.57)	(0.53)		(0.10)	(0.13)		(0.66)	(0.96)
Treatment*% of treatment households within 1 km			0.01			0.00			0.00
			(0.10)			(0.61)			(0.58)
Number of total households within 1 km		0.00	0.00		0.00	0.00		0.00	0.00
		(0.33)	(0.31)		(0.54)	(0.55)		(0.80)	(0.78)
Treatment saturation Min (out 100%)	18.18	18.18	18.18	18.18	18.18	18.18	18.18	18.18	18.18
Treatment saturation Max (out 100%)	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00
Treatment saturation Mean (out 100%)	50.37	50.37	50.37	50.37	50.37	50.37	50.37	50.37	50.37
Panel B: Radius of 3km									
Treatment	0.04	0.04	-1.24	0.02	0.02	0.81	0.01	0.00	-0.07
	(0.36)	(0.38)	(0.14)	(0.66)	(0.67)	(0.34)	(0.70)	(0.76)	(0.80)
% of treatment households within 3 km		0.01	-0.01		0.00	0.01		0.01*	0.00
		(0.48)	(0.64)		(0.71)	(0.37)		(0.06)	(0.26)
Treatment*% of treatment households within 3 km			0.03			-0.02			0.00
			(0.12)			(0.35)			(0.79)
Number of total households within 3 km		0.00	0.00		0.00	0.00		-0.00***	-0.00***
		(0.82)	(0.82)		(0.59)	(0.59)		0.00	0.00
Treatment saturation Min (out 100%)	35.48	35.48	35.48	35.48	35.48	35.48	35.48	35.48	35.48
Treatment saturation Max (out 100%)	57.39	57.39	57.39	57.39	57.39	57.39	57.39	57.39	57.39
Treatment saturation Mean (out 100%)	49.93	49.93	49.93	49.93	49.93	49.93	49.93	49.93	49.93
Panel C: Radius of 5km									
Treatment	0.04	0.04	-1.52	0.02	0.02	1.74	0.01	0.00	-0.31
	(0.36)	(0.37)	(0.21)	(0.66)	(0.70)	(0.15)	(0.70)	(0.72)	(0.36)
% of treatment households within 5 km		0.00	-0.02		0.00	0.01		0.00	0.00
		(0.99)	(0.42)		(0.86)	(0.46)		(0.80)	(0.70)
Treatment*% of treatment households within 5 km			0.03			-0.03			0.01
			(0.20)			(0.15)			(0.36)
Number of total households within 5 km		0.00	0.00		0.00	0.00		-0.00**	-0.00**
		(0.84)	(0.84)		(0.72)	(0.73)		(0.05)	(0.05)
Treatment saturation Min (out 100%)	43.97	43.97	43.97	43.97	43.97	43.97	43.97	43.97	43.97
Treatment saturation Max (out 100%)	52.93	52.93	52.93	52.93	52.93	52.93	52.93	52.93	52.93
Treatment saturation Mean (out 100%)	50.22	50.22	50.22	50.22	50.22	50.22	50.22	50.22	50.22
Observations	1980	1980	1980	1953	1953	1953	1968	1968	1968
Control Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.90	0.90
Control SD	1.00	1.00	1.00	1.00	1.00	1.00	0.31	0.31	0.31

Notes: Pvalues are in parenthesis. The regressions include baseline controls and baseline values of the dependent variable, where applicable

Appendix C: Timeline

While recruitment of GiveDirectly cash transfer recipients began in March, the first wave of participants started receiving sequence 1 of behavioural nudges on 2nd August. On 11th October, wave 2 participants joined the intervention by receiving their first sms of sequence 6 nudges along with wave 1 participants. The third wave of the intervention began receiving their first message (sequence 8 of the nudges) on 8 November. Altogether, Wave 1 had 23 weeks of the intervention while wave 2 and 3 had 13 and 9 weeks of treatment, respectively. **Figure C1** shows the overall evaluation timeline.

Figure C1: Overall evaluation timeline

Appendix D: Child protection protocol

D.1 Reach protocols

Enumerators will schedule calls with parents/guardians to ensure both parents/guardians and children are free for the interviews. Once calls have been scheduled, Enumerators will send SMS reminders to parents/guardians the night before and the morning of the call. Enumerators will then try to call parents/guardians at the agreed-upon time to conduct the interview. If the parents/guardians or their children are unavailable, Enumerators will reschedule the call up to 7 times.

Reminder SMS the night before the call:

Hello [parent_name],

This is a reminder that you have an interview scheduled for tomorrow at [the agreed upon time] with [Enumerator name]. We would appreciate it if you could make sure that you and [name of child] are available during this time for the interview.

Reminder SMS on the morning of the call:

Good morning [parent_name],

This is a reminder that you have an interview scheduled for today at [the agreed upon time] with [Enumerator name]. We look forward to talking with you and [name of child].

D.2 Order of the interview section

Enumerators will start by building rapport with parents by asking them questions about the parent/guardian's highest level of education, their aspirations for children, how their child spends their time, attendance, and academic progression. The Enumerator will then ask the parent to hand the phone over to the child and begin the child interview. The child interview should be at most 15 minutes. Once the Enumerator has completed the child interview, the Enumerator will ask the child to hand the phone back to the parent for the rest of the parent interview, which should last roughly 20 minutes. The whole interview process, including the consent reading, should not exceed 40 minutes.

The order of the interview questions for the pilot will be as follows:

- Parents:
    - Consent for self and child
    - Education level of parents
    - Aspirations for children

- - - Children's time spending (yesterday)
    - Attendance (past 1 week)
    - Academic progression
    - Receipt/understanding of SMS message
  - [Passing the phone to children]
  - Children:
    - Rapport building (asking about the favourite dish, game, what school they go to, what grade they are in and what they like to do when they are not in school)
    - General engagement with parent plus attendance yesterday
    - Engagement of parent with school life
    - Time spending (yesterday)
    - School motivation
  - Parents:
    - General engagement with child
    - Engagement with child's school life
    - Parental Child Discipline practices

D.3 Obtaining consent from parent

When the Enumerator calls and both the parent/guardian and child are available, the Enumerator will proceed to read out the Study Information Sheet to the parent. If the parent has any questions, the Enumerator will address them and then seek verbal consent for their participation as well as their child's participation before the interview begins. Should parents choose to refuse their participation, then that marks the end of the interview. If they choose to participate but choose not to provide permission for their child to participate, the interview will still be conducted with only the parent.

Consent for parents:

My name is ${enum_name}, and I am assisting IDinsight, a non-government organisation (NGO), with a research project. We are working with GiveDirectly and Movva (an education NGO) on a research project to understand parental engagement with children's education. We received this phone number through an interview that GiveDirectly conducted with ${main_recipient} who registered this as the phone number to reach your household. As a part of this research, we would like to speak with you and your child ${name}.

Introduction and Purpose of Research
The decision for you and your child to participate or not participate in this research will not affect your chances of participating in any future or present GiveDirectly programs. This consent form may contain words that you do not understand. Please ask me to stop as we go through the information, and I will take the time to explain. If you have questions later, you can ask them to me or to another researcher.

Participant Selection and Procedures
As a part of this study, we are selecting GiveDirectly recipients with children of school-going age to do a 1-time survey. The survey will take not more than 30-45 minutes. We will first ask you some questions about how you engage with your children, and then we would like to ask you to pass the phone to one of your children (the child has been selected from the list of child names

that you provided to GiveDirectly). After that, we will ask your child questions on similar topics. The information given is confidential, and no one else except the research team will have access to the information that is tied to your name.

Voluntary Participation and Right to Refuse or Withdraw

You and your child's participation in this research is voluntary. You and your child may decline to participate in the survey or choose not to answer specific questions. Refusal to take part will not result in any penalty or loss of benefits to which you are otherwise entitled. If you and your child decide to participate, you are free to withdraw at any time without penalty or loss of services, and without providing any reason for your withdrawal. If there is anything that you or your child would prefer not to discuss, please feel free to say so. If you and your child choose to participate in this research study, we will require verbal consent from both you and your child before we proceed with the interview. If you or your child wish to withdraw any information you provided during the survey, don't hesitate to contact IDinsight.

Confidentiality and Risks

We will make every effort to keep all data strictly confidential, from how it is stored to who and how it is shared.

Compensation and Benefits

You and your child will not be compensated for participating in this survey. Therefore, there are no direct and immediate benefits to you or your child from participating in this survey. Still, we hope that our results will help better serve future program participants with children.

Sharing the Results

The knowledge we get from this research will be shared with GiveDirectly and Movva before it is widely available to the public. We may publish the results, but without information that can identify you, so that other interested parties may learn from the research.

Who to Contact

This study has been reviewed and approved by the National Committee on Research in the Social Sciences (NCRSH), a committee responsible for ensuring that research participants are protected from harm. If you have any questions relating to your rights or seek to report a violation of your rights, please contact Malawi National Committee on Research in Social Sciences via phone at +265 771 550/774869 or email at ncrsh@ncst.mw.

If you have any questions, you can ask them or use the contact information below.

We will send you contact information in case you have any questions or concerns via SMS.

"Hello [recipient name],

My name is [enumerator name] and I called earlier today about the IDinsight research project. For any concerns or questions you may have for IDinsight following the interview please contact our Country Lead T. Arthur Chibwana on +265 884 715 707. For any concerns or questions you may

have relating to your rights or if you wish to report a violation of your rights, please contact Malawi National Committee on Research in Social Sciences via phone at +265 771 550/774869."

D.4 Explaining to parents why we want to talk to children

After Enumerators have obtained consent from parents, they will explain to parents why they need to talk to their children and where they should be when their children are being interviewed. Enumerators will highlight that:

- the research study wants to understand how children spend their time and how they interact with other household members
- the conversation with their child is of low stakes and does not ask any sensitive questions that they should worry about
- their child can choose not to participate entirely or answer any question they feel uncomfortable answering
- parents should provide their child with a quiet space to respond to the study questions
- they should not be present, or the call be on speakerphone, so the child feels free to talk to the Enumerator and to maintain privacy

Enumerators can use the below script:

"Hello! My name is [...] and I am looking to speak with [name of parent] and your child [name of child] if they are available."

…

"Now I'd like to speak to [name of child]. I will ask [name of child] questions about how their parents/guardians relate with them, how they spend their time, and what motivates them to go to school. The interview will take about 15 minutes. During this interview, [name of child] may not answer any questions if they do not want to, and I will do my best to make them comfortable. You may pass the phone to [name of child] and be close to the conversation, but I would like [name of child] to answer independently, describing their experience. The questions will be multiple-choice, and [name of child] will be asked to answer whether something happens "always, sometimes, never" to a set of questions. Could you please pass the phone to [name of child]?"

D.5 Instructions what to tell parents when they are passing the phone

Once parents have understood and agreed for the Enumerators to interview their child, the Enumerator will ask the parent to pass the phone to the child. Before parents hand the phone over to the child, Enumerators should instruct parents to briefly explain why the Enumerator wants to talk to them and assure them that it is safe to do so. For example, Enumerators will ask parents to say the following:

"I just talked to ${Enumerator_name} who is trying to learn how different parents/children like us spend time together. They have already talked to me, and now they would like to hear from you. Remember that if you don't understand what they are asking, you can tell them so. I won't listen to the call, but I'll be nearby."

D.6 Practices to adopt when talking to children

When speaking to children, enumerators will make sure to create a safe environment where children feel cared for and safe. In addition, Enumerators will use encouraging and supportive language to help children fully express themselves. They will respect what the children say, be optimistic about their responses, refrain from judging or correcting them, and encourage them to respond authentically. Enumerator conversations with children will also be monitored using SurveyCTO and audio checks.

Enumerators will follow the guidelines below for talking to children:

- **Relaxed environment:** Enumerators will ensure that the interview does not feel like an interrogation. They will encourage children to relax and feel free. Before they start the interviews, they will have a friendly conversation with the child asking the questions like, "what's your favourite dish?", "what's your favourite game to play?", "what do you enjoy doing when you are not in school?". These questions will help the child feel more relaxed. Enumerators will tell and assure the child that whatever they say will not directly be shared with anyone, including their guardian. They will also say to the child that they can omit any questions they don't want to answer and will be allowed to do so.
- **Quiet space, no distractions:** Enumerators will tell the child to find a quiet room, far from distractions. If other children gather around, Enumerators will tell the child to ask them to leave. Before starting the interview, the Enumerator will ask the child where they are currently located (outside/inside) and if they feel comfortable answering the question where they are.
- **Encouragement and patience:** Enumerators will listen carefully and encourage the children by appreciating their effort and participation. Enumerators will give children enough time to understand what is being asked, think, and respond to questions. Enumerators will be patient and not hurry the children. They will repeat or rephrase questions so that the children fully understand them.
- **Child protection:** Before starting your work with a child, do an 'ice-breaker' to help them relax and feel comfortable. Make sure that children know that they are allowed to stop the survey at any time. Use language that is appropriate, respectful, and clear, (but do not change any words on the survey! If any of the language there is difficult for children to understand, let your field manager know right away. Avoid asking a child about his or her personal or home life.
- In case an Enumerator has concerns about a child's well-being, signs of abuse, or other child protection issues, they should immediately report what they've noticed to the IDinsight Field Managers for help in reporting the case to relevant authorities like Social Welfare Services and Victim Support Units.

D.7 Interviewing children

Enumerators will start by calling and obtaining consent from parents/guardians to interview school-age children in their households. If the parent/guardian consents, they would be asked to pass the phone to the child, giving them room to talk freely. Enumerators will also clarify that the questions are 'low stakes' and have no rewards and consequences associated to help with facilitating honest responses.

Enumerators will use the following introductory statement will to introduce the interview and obtain assent from younger children: *"Hello! My name is [...], and I am here today to quickly chat about how you and your family spend time together! I just talked to your parent/guardian X, and now I want to hear from you. There are no right answers, and I am curious to understand your experience. What you say will not be shared with anybody. If you don't want to answer a question, say, "I don't want to*

answer this". Please find a quiet space where you can hear me well. Are you okay talking to me for a few minutes right now?"

In the event that Enumerators overhear anyone helping the child answer, they will read out the following statement: *"I hear somebody else's voice in the background. I would like to hear what you think, not somebody else. Can you please find a space where you feel comfortable answering questions? If not, then, I can talk to your parent/guardian again. Are you comfortable now?"* If the child says yes, continue with questions.

If a child sounds very nervous, is unwilling to talk, or not paying attention, Enumerators will ask the parent/guardian for support and consider giving the child more time or another chance at a later point during that day or another day. Children often know and trust their parents/guardians; having their reassurance will help them feel more comfortable talking to our team. If a child refuses to participate even with the support of his or her parent/guardian, Enumerators will indicate that the child has not consented to be interviewed and continue with the rest of the interview.

Appendix E: Heterogeneous treatment effects

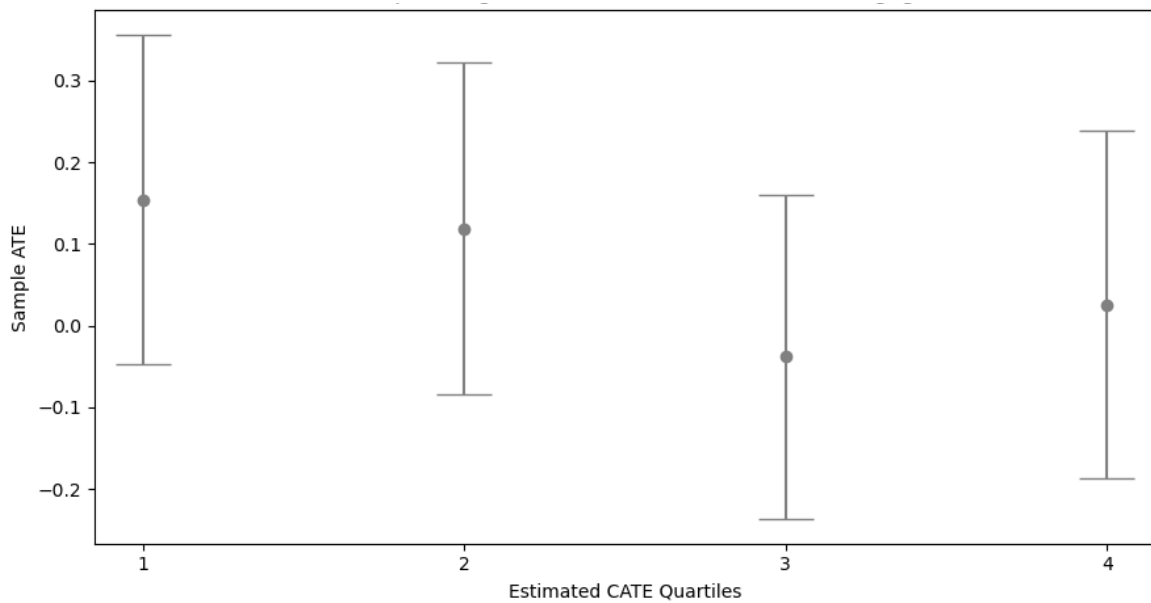**Figure E1**: Sorted group average treatment effects for outcomes

Figure E2: Sorted group average treatment effects for child reports of attendance
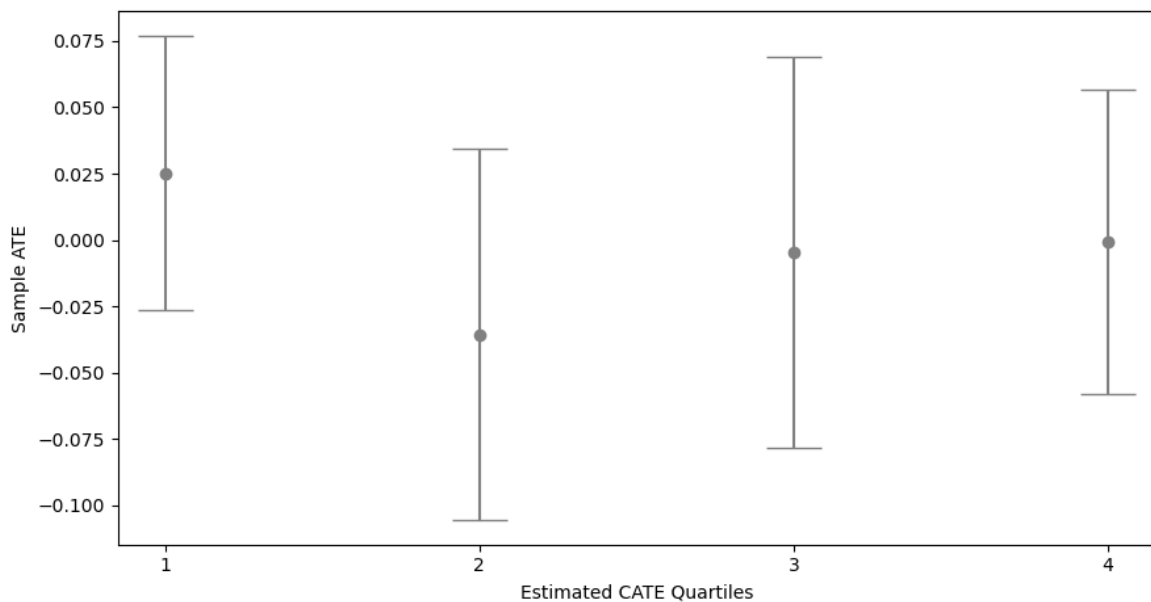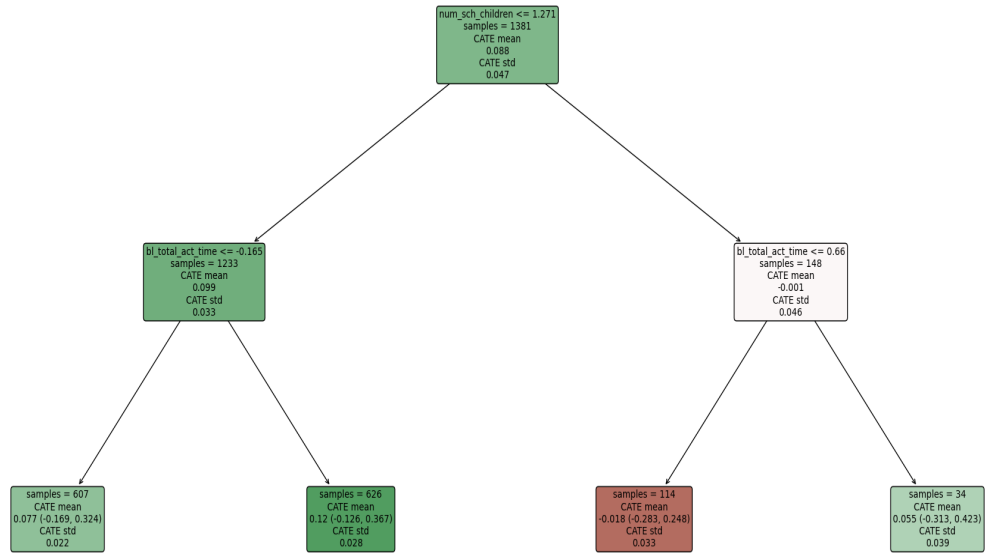
Figure E3: Causal tree for parental engagement