Education Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cede20

Using SMS and parental outreach to improve early reading skills in Zambia

Alejandro Ome & Alicia Menendez

To cite this article: Alejandro Ome & Alicia Menendez (2022) Using SMS and parental outreach to improve early reading skills in Zambia, Education Economics, 30:4, 384-398, DOI: 10.1080/09645292.2021.1988518

To link to this article: https://doi.org/10.1080/09645292.2021.1988518

Published online: 12 Oct 2021.

EDUCATION ECONOMICS
2022, VOL. 30, NO. 4, 384–398
https://doi.org/10.1080/09645292.2021.1988518

Routledge
Taylor & Francis Group

Using SMS and parental outreach to improve early reading skills in Zambia

Alejandro Ome[a] and Alicia Menendez[b]

[a]NORC at the University of Chicago, Chicago, IL, USA; [b]Harris School of Public Policy, The University of Chicago and NORC at the University of Chicago, Chicago, IL, USA

ABSTRACT
We evaluate an intervention using SMS messaging to send short stories to 2nd and 3rd graders in Zambia's Eastern province, aimed at improving reading skills. For nine months households received three text messages weekly comprising a short story for children to read with their families, and a question about the story. Additionally, parents attended monthly meetings aimed to address any program implementation issues and encourage reading. The program had a positive impact on reading skills, between 19 and 28 percent of a standard deviation. A cost-effectiveness analysis shows that expanding the program nationally would cost USD 20–22 per child.

ARTICLE HISTORY
Received 9 December 2020
Accepted 15 September 2021

KEYWORDS
Human capital; economic development; SMS; early reading; Zambia

JEL CODES
C93; D91; I25

1. Introduction

Using information and communication technologies (ICT) constitutes a potentially powerful channel to disseminate access to educational resources. This is particularly relevant in developing countries where, on one hand, access to printed materials is restricted due to economic hardships, and on the other, access to ICT devices has been increasing dramatically, in particular cell phone use.[1] In this study we evaluate the impact of the *Makhalidwe Athu*[2] project (MA), an intervention that used SMS messaging to send short stories to roughly 1,200 2nd and 3rd graders in Zambia's Eastern province, aimed at improving their reading skills. The SMS served a dual purpose. First, it provided households with age appropriate and culturally relevant reading materials, which are generally scarce in these contexts. Second, the SMS acted as a reminder for parents to read with children, or at least give children time to read.[3] In addition, participant parents attended monthly meetings to address any program implementation issues and encourage reading.

A large literature documents the impact of using computers and other ICT devices in schools (McEwan (2015) and Bulman and Fairlie (2016) review these studies). The impact of using ICT devices at home has received less attention, but the causal studies that have been conducted, most of them looking at the impact of computers, usually fail to find positive impacts on academic achievement (Malamud and Pop-Eleches 2011; Beuermann et al. 2015; Bulman and Fairlie 2016; Cristia et al. 2017). Although less than 15 percent of the population in Zambia has access to the internet, we estimate that at least 60 percent of the population has access to a cell phone. Access to computers is obviously not comparable with SMS messaging; for starters, the volume of educational resources that can be channeled through the most basic computer dwarfs anything that can be transmitted through SMS; however, one feature of MA that is worth highlighting that other

CONTACT Alicia Menendez menendez@uchicago.edu Harris School of Public Policy, The University of Chicago and NORC at the University of Chicago, 1307 E 60th St, Chicago, IL 60637, USA
Supplemental data for this article can be accessed https://doi.org/10.1080/09645292.2021.1988518.

home-based ICT interventions lack is how parental participation was harnessed. In effect, MA involved a behavioral component targeted at parents, not only via the reminder nature of the SMS, but through the monthly meetings with program facilitators.

This study contributes to the growing literature on ICT devices readily owned and used by the target population, specifically cell phones, to deliver reading materials and encourage reading at home. Examples of SMS-based concepts that have been used in education in developing countries include delivering quizzes and mini-lessons to students in Kenya (Eneza education, Reid and Pruijsenb 2015), promoting communication and literacy through an SMS-based forum in Senegal (Jokko initiative, Hanemann and Scarpino 2016), targeting adult education classes (Aker, Ksoll, and Lybbert 2012), and disseminating teaching resources for teachers in Papua New Guinea (SMS Story, Kaleebu et al. 2017). MobiLiteracy-Uganda is a program similar in spirit to MA but smaller both in terms of program duration and content. For three months, participant parents received SMS and audio messages in their phones introducing all the letters of the Luganda alphabet, and 10 key vocabulary words. Similarly, to MA, the objective was that parents read these SMS with their children. To evaluate the program on children's reading skills, Pouezevara and King (2014) used a sample of about 55 households (parents and their children) per study group and found positive but modest impacts on reading skills. In the US, recent studies have also found positive impacts of programs that use text messaging as reminders for parents to support their children's development (Mayer et al. 2018; York, Loeb, and Doss 2019).

We find that MA had a positive impact on reading skills between 19 and 28 percent of a standard deviation depending on the measure used. We also explore the channels through which the program had this impact, namely if children were reading more at home. We show that MA had a positive impact on the amount of time children read on their own at home. We also find some evidence that students read more at home with their parents, although this result is not robust to different specifications.

Finally, we estimate the per-student cost of scaling up the program at the national level and find that this will cost between USD 20–22 per student.

This paper has five sections including this introduction. We describe the program in the second section. The third section presents the data and evaluation design. In the fourth section, we discuss the evaluation results and the costs of the intervention. The fifth section concludes.

2. Project background

Zambia is a landlocked country in Southern Africa. It ranks 131 out of 157 countries in the Human Capital Index, a construct that combines health and education indicators to approximate the amount of human capital that children born today will gain by age 18 (World Bank 2018b). As other developing countries, in the last couple of decades, Zambia observed a substantial increase in primary school attendance. However, multiple challenges have prevented the education imparted from having acceptable quality. This is apparent when looking at early grade reading performance. According to a reading assessment conducted in 2014 on a national sample of 2nd graders, between 46 and 88 percent of students – depending on the local language- were not able to read a single word (Brombacher et al. 2015).[4]

Lack of reading materials and exposure to print are potentially some of the factors driving these poor reading results. Baseline data collected for this study show that only 58 percent of students have access to reading materials outside of school. The main objective of MA was to provide reading materials for children to read at home.

Creative Associates designed and implemented the program, which was funded by All Children Reading Partners (USAID, World Vision and the Australian Government). Between April and December 2016 MA participant households received three text messages on their mobile phones each week, transmitted on Mondays, Wednesdays and Fridays. Each of these three messages (with a maximum of 160 characters each) comprised part of a short weekly story for children to read with

their families, as well as a question about the story. Participants could also call in for a pre-paid recorded voice message, which included comprehension questions, as well as a recording of the story itself.

Most stories were submitted by community members that learned about the program through local media and community meetings. MA's local reading specialist conducted workshops with the teachers and staff from the Ministry of Education so the submitted stories were leveled for Grade 2 and 3 readers and adapted to fit the SMS format. A total of 41 stories were transmitted between March 2016 and Feb 2017. We reproduce one of these stories in Annex I.

The underlying theory of change of the program was that by providing reading materials that were culturally relevant and entertaining, children were likely to read more and become better readers.

Following an initial orientation meeting with caregivers, head teachers and students about the program, community mobilizers met monthly with caregivers. During the orientation meeting, community mobilizers explained the purpose of the short SMS stories and their structure, discussed the need for caregivers to set time to listen to the child read the story and answer the follow-up questions, and conducted a demonstration on how to access the stories. Monthly meetings were devoted to encouraging community participation in the SMS story development, discuss and address any issues that had arisen in the context of the program, and emphasize the need for the student to be the one reading the stories. The program also provided children with a notebook and encouraged them to write the stories.

An uptake survey conducted on a subsample of caregivers in the treatment group in June 2016 showed that participation in the program was high. Ninety-five percent of respondents said that they had attended at least one MA meeting. Ninety-two households reported they were receiving the MA SMS and of these, all caregivers said that they read the messages with the child at least once a week. The endline data reflected a high level of program uptake as well. Ninety-three percent of caregivers in the treatment group reported they had attended at least one MA meeting or training session; these respondents attended an average of seven meetings over the life of the program. Eighty-four percent of caregivers in the treatment group said they received stories via SMS. Eighty-one percent of caregivers in the treatment group showed the text messages to the child at least once a week. Participation was also high according to the children's reports. Eighty-four percent of children in the treatment group said that they have received SMS stories, and 81 percent of children said they had read the last story they received.

3. Data and evaluation design

To construct the sample, we randomly selected 80 schools from a list of 293 public schools eligible to participate in two districts of Eastern Zambia, Chipata and Lundazi.[5] We randomly assigned half the schools to the treatment group and half to the control group.[6]

In October 2015, we screened all students in grades 1 and 2 across all 80 sampled schools to determine whether anyone in their home had a cell phone.[7,8] In total, enumerators attempted screening 8,681 students during school visits. This includes 2,354 students (27.1 percent) that were absent during the visits. Of the students attending on screening day, 4,910 (56.6 percent of the total) reported having a cell phone at home. In each school, we sampled 30 of the students that reported having a cell phone at home.

The caregivers of the sampled students were interviewed between November and December 2015. The caregiver interview included a series of screening questions intended to determine the respondent's eligibility to participate in the study. This included questions about the respondents' interest in participating in the intervention and the respondents' (actual) access to a working cell phone number. The same screening process was conducted in both treatment and control groups. Respondents did not know to which group they were going to be assigned when we interviewed them. In the cases that multiple children from the same household were sampled,

household-level questions were only asked once, but student-specific questions were asked for each child. We completed 2,222 caregiver interviews that correspond to 2,397 students.[9] Table 1 summarizes the number of students and parents surveyed at each wave of data collection.

In January 2016, enumerators visited schools to survey the sampled students. To assess students reading skills we used a version of the Early Grade Reading Assessment – EGRA (RTI International 2015) created for Zambia. EGRA is a widely used test to assess early grade reading performance. The instrument we used contains 5 subtasks in ChiNyanja language, namely letter sound identification, non-word decoding, oral reading fluency, reading comprehension, and listening comprehension. Appended to the EGRA was a short student questionnaire to capture children's reading habits at home and other reading-related information. In total 2,263 students were assessed and interviewed. This is lower than the number of students for which we collected student-level caregiver data (2,397) and was mostly due to student absenteeism the days EGRA data was collected. Importantly, enumerators were not able to collect EGRA data in one of the 80 schools because weather conditions made it inaccessible.

Endline data was collected between mid-January and early February of 2017. Instruments administered at endline were the caregiver questionnaire, student questionnaire, and the EGRA. For purposes of comparability, the same EGRA tool was used at baseline and endline.

In total 1,942 caregivers were surveyed at baseline and endline, that correspond to 2,091 children (there are more students than caregivers because we have 149 siblings in the sample). Out of these 2,091 children for whom we have caregiver data, 2,054 were assessed and interviewed at endline. The longitudinal sample of students assessed with EGRA both at baseline and endline is 1,973. This is the sample that we use for the analysis in this paper. The attrition rate for the parent survey was 12.6 percent and for the student survey was 12.8 percent. Attrition was not correlated with treatment status (the *p*-value for the student survey is 0.61 and for the parent survey is 0.85).

Table 2 shows students' and households characteristics and students' reading scores, by treatment status. There are no significant differences in students' characteristics (Panel A) between treatment and control groups. We show household characteristics in Panel B. The differences between treatment and control groups are small and not statistically significant at standard levels of confidence. The last variable in Panel B, the asset index, corresponds to the first principal component of the household ownership of the following assets: chair, bed, clock, radio, television, computer, bicycle, motorcycle, car, fridge and stove.

Panel C shows baseline statistics on reading habits and attitudes towards reading. We asked caregivers and children how often the latter read at home with someone else and on their own. According to the caregivers, on average, 39 percent of them never read with the MA child, and among the ones that read with children, the majority read between two and three times per week (roughly a third of all caregivers).[10] Instead of asking the same categorical question we asked parents, to make it simpler for children we asked them how many days they had read with someone the previous week. The results indicate that, on average, children read with someone at home 2.3 days per week. Regarding reading on their own at home, according to the caregivers about 43 percent of children never read on their own at home, and within the ones that do read on their own, the majority read between two and three times per week (roughly a fourth of the total). According to children's own reports, they read on their own at home 2.4 days per week, on average. In terms of balance, we only found a statistically significant difference for reading with the caregiver, as

Table 1. Sample.

	Target	Baseline	Endline	Panel	Attrition
Caregiver survey					
Caregiver level data	2,400	2,222	1,942	1,942	12.6%
Student level data	2,400	2,397	2,091	2,091	12.8%
Student survey	2,400	2,263	2,054	1,973	12.8%

Table 2. Descriptive statistics at baseline.

Variable	Control		Treatment		Diff	
A. Student characteristics						
Female	0.52	(0.02)	0.51	(0.02)	0.02	(0.02)
Age	8.79	(0.07)	8.85	(0.06)	−0.05	(0.1)
Third grade	0.48	(0.01)	0.48	(0.01)	−0.01	(0.01)
B. Household characteristics						
Caregiver can read	0.71	(0.02)	0.72	(0.02)	−0.02	(0.03)
Household size	5.37	(0.13)	5.23	(0.13)	0.14	(0.18)
There is at least one children's book at home	0.44	(0.02)	0.46	(0.02)	−0.02	(0.03)
Asset index	−0.01	(0.08)	0.01	(0.07)	−0.02	(0.11)
C. Reading habits						
Reading with someone at home						
- According to the caregiver[1]					−0.14*	(0.08)
Never	0.36		0.41		−0.05	
Less than once a month	0.02		0.01		0.01	
Once or twice a month	0.04		0.06		−0.02	
One day a week	0.15		0.14		0.01	
Two or three days a week	0.33		0.31		0.02	
Four days a week or more	0.11		0.08		0.03	
- According to the child (days per week)	2.27	(0.10)	2.37	(0.10)	−0.10	(0.14)
Reading alone						
- According to the caregiver[1]					0.11	(0.08)
Never	0.42		0.45		−0.04	
Less than once a month	0.02		0.01		0.00	
Once or twice a month	0.03		0.03		0.00	
One day a week	0.12		0.13		−0.01	
Two or three days a week	0.25		0.25		0.00	
Four days a week or more	0.16		0.13		0.04	
- According to the child (days per week)	2.33	(0.13)	2.46	(0.10)	−0.14	(0.16)
D. Attitudes towards reading						
Child likes to listen to stories	0.57	(0.02)	0.54	(0.02)	0.03	(0.03)
Child likes to read or practice reading	0.63	(0.02)	0.61	(0.02)	0.02	(0.03)
E. Reading skills						
PCA reading index	0.00	(0.08)	0.19	(0.07)	−0.19*	(0.11)
Oral reading fluency (wpm)	3.76	(0.63)	5.09	(0.57)	−1.33	(0.84)
Reads at least one word	0.29	(0.03)	0.36	(0.03)	−0.06	(0.04)

Notes: The asset index is the first principal component of the household ownership of the following assets: chair, bed, clock, radio, television, computer, bicycle, motorcycle, car, fridge and stove. The PCA reading index is the first principal component of the five EGRA scores, standardized using the control mean and standard deviation. Sample sizes are lower than 1,973 for some outcomes due to item-specific missing data.

[1]To estimate whether there are significant differences between treatment and control for ordered categorical variables (caregivers' responses on how often they read with the child and how often the child reads alone) we estimated ordered probit regressions where the outcome of interest are the ordered categorical variables and the covariate is the treatment dummy; statistical significance is determined by the test on the treatment dummy, hence in the third column we show the coefficient and standard error on the treatment dummy. Standard errors clustered at the school level in parentheses.

*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

reported by the caregiver, suggesting that caregivers in the control group read with children more often than their counterparts in the treatment group.[11]

Panel D shows results on attitudes towards reading. We asked children whether they enjoy listening to stories and if they enjoy reading. We provided them with three response options: 'Like', 'Indifferent', and 'Do not like'. We recoded the responses to create an indicator variable that is 1 if the child answered 'Like' and 0 otherwise. At baseline a little more than half the children reported they like listening to stories, and almost two-thirds said they like reading or practice reading.

Panel E shows baseline reading skills. We focus on three indicators. The first one is the PCA reading index; this is the first principal component of the five EGRA scores, standardized using the control mean and standard deviation, pooling both 2nd and 3rd grades.[12] This is our preferred outcome of interest because it summarizes all the results of the reading assessment. We also show results for two indicators derived from the oral reading subtask. The first is the number of words correctly read from a short passage in one minute, commonly called oral reading fluency (ORF).

This indicator is widely used to benchmark early reading proficiency. In the case of Zambia, official benchmarks indicate that 2nd graders should read a minimum of 25 correct words per minute (cwpm), while 3rd graders should read a minimum of 40 cwpm (Ministry of Education, Science, Vocational Training and Early Education 2014). Notably, children in this sample were reading only 4.4 cwpm at baseline, on average, which highlights how behind these children are relative to the official benchmarks.[13] Finally, our third and last indicator is the fraction of students that are able to read at least one word correctly in the oral reading fluency subtask. We report this indicator given the large fraction of second and third grade students that are not able to read a single word in this sample and across the country.

While the differences at baseline between treatment and control groups are significant only for the PCA reading index, all three reading skills scores are slightly higher for the treatment group than for the control group despite treatment status was randomized. To isolate these baseline differences, we include baseline reading scores as covariates in our estimation of the effect of the program. This approach mimics the canonical value-added model, which has been found to replicate experimental parameters in education contexts (Abdulkadiroglu et al. 2011; Deming 2014; Deutsch 2012; Singh 2015). We also include as covariates households' and students' characteristics collected at baseline. The regression model we use to estimate the impact of the program can be described by:

$$A_{s1} = \alpha + \beta D_s + \gamma A_{s0} + x_s{}'\delta + u_{st} \tag{1}$$

where, A_{s1} and A_{s0} measure achievement (EGRA score) for student s at endline and baseline, respectively; D_s is a dummy variable indicating treatment status; $\mathbf{x_s}$ is a vector of characteristics of the student (age and its square, gender and grade) and of the student's household (caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, the household asset index, and an indicator variable for whether ChiNyanja is the language most commonly used at home); u_{st} is an error term and α, β and $\boldsymbol{\delta}$ are parameters to be estimated. The coefficient of interest is β, which measures the contribution of the intervention to student achievement.

Given the imbalance in the outcomes of interests at baseline, as robustness checks we estimate the impact of the program using three additional methodologies. First, we estimate a standard difference-in-difference model, where the outcome of interest in Equation (1) is not the endline value but the difference between baseline and endline of the three outcomes of interest we focus on (and the lagged value of the outcome of interest is not included as a covariate). This is equivalent to the value-added model in Equation (1) except that it assumes no depreciation in reading skills ($\gamma = 1$).

Second, we use matching methods to provide a more flexible way to control for the differences at baseline between treatment and control, specifically we use entropy matching to produce weights that minimize differences in the distributions of students' characteristics at baseline (Hainmueller 2012).

The covariates included in the matching and the difference-in-difference regressions are the same variables included in Equation (1). We collected a number of additional sociodemographic characteristics at baseline. In our third robustness check, we use a Lasso regression to expand the number of covariates as potential predictors. These additional covariates are the five EGRA subtasks scores, district fixed effects, language most commonly used at home dummies, number of children books in ChiNyanja, four variables to account for household composition by age (0–5-year-olds, 6–15-year-olds, 16–59-year-olds, and 60-year-olds and older), an indicator for whether the house's floor is made of mud or dirt, self-reported child's school attendance in the previous week, and an indicator variable for being the first born. To produce Lasso results for inference we use the double-selection algorithm proposed by Belloni, Chernozhukov, and Hansen (2014).[14]

4. Results

4.1. Reading skills

Table 3 shows the impact estimates of MA on reading skills. In the first column, no covariates are included. The PCA reading index is standardized using the mean and standard deviation of the control group at each wave of data collection, so the coefficients can be interpreted directly as percentages of a standard deviation. The difference in the PCA reading index at endline between treatment and control groups is 33 percent of a standard deviation. When the baseline value of the PCA reading index is included as a covariate, as we show in column (2), the difference between treatment and control groups falls to 20 percent of a standard deviation. This change shows the magnitude of the imbalance between the two experimental groups at baseline. Adjusting for the baseline value seems to correct it. Including additional covariates in the regressions -as we do in column (3)- does not change the coefficients in a meaningful way. We show the estimated for oral reading fluency in the second row of Table 3. The unadjusted difference between treatment and control groups indicates a program impact of 4.66 cwpm, but the coefficient becomes 3.46 cwpm when the baseline value of the dependent variable is included as a covariate, and 3.35 cwpm (0.28 of a standard deviation of the endline scores for the control group) when all the covariates are included.[15] To put these effects in perspective, for oral reading fluency the treatment group shows twice as much improvement as the control group over the same time period. Gains in oral reading fluency for the control group amount to 3.7 cwpm between baseline and endline, while the equivalent gains in the treatment group are 7 cwpm. In other words, the MA program gains are of similar magnitude to what 2nd and 3rd graders (on average) gain in a full year of regular schooling.

The impact of MA compares well with that of other interventions aimed at improving early grade reading, particularly if we take into account that MA is a low intensity intervention that does not require resources from the education system. Graham and Kelly (2018) summarize the impact of 18 early grade reading interventions in less developed countries. All the evaluations used experimental or quasi-experimental methodologies, target the same grades, and use EGRA to measure reading performance. They found effect sizes in oral reading fluency that range from −0.02–0.80 of a standard deviation, with a mean of 0.38 and a median of 0.36 of a standard deviation. Murray et al. (2021) report that the effect sizes associated with four evaluations of early grade reading interventions in Latin America range from zero to 0.32 of a standard deviation.

Table 3. Impact of MA on reading skills.

	(1)	(2)	(3)
PCA reading index	0.33**	0.20***	0.19***
	(0.10)	(0.050)	(0.049)
	[0]	[0]	[0]
Oral reading fluency (wpm)	4.66***	3.46***	3.35***
	(1.14)	(0.73)	(0.73)
	[7.4]	[7.4]	[7.4]
Reads at least one word	0.14**	0.097***	0.096***
	(0.041)	(0.026)	(0.026)
	[0.44]	[0.44]	[0.44]
Baseline value of the dependent variable included as covariate		X	X
Sociodemographic controls			X
N	1968	1968	1968

Note: Regressions on PCA reading index have five fewer observations due to item specific missing data in EGRA subskills other than ORF. Outlier values, one for oral reading fluency and two for non-word decoding, both at baseline, are set to missing. Sociodemographic controls include student's age and its square, gender and grade, caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, and an indicator variable for whether ChiNyanja is the language most commonly used at home and the household asset index. Standard errors clustered at the school level in parentheses. Mean value of the dependent variable at endline for the control group in squared brackets.

*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

Finally, we estimate the program effect on an indicator variable for whether students are able to read at least one word from a grade leveled passage and show the results in the third row of Table 3. The unadjusted difference between treatment and control at endline is 14 percentage points. When the baseline value of the dependent variables is included, the coefficient becomes 9.7 percentage points, and it does not change substantially when the rest of the covariates are included.[16]

The imbalance in reading skills between treatment and control groups at baseline is a source of concern for identification. As a robustness check Table 4 shows impact results for the three outcomes of interest using difference-in-differences, entropy matching and LASSO regression.

The difference-in-difference model is similar to the value–added regression except that the coefficient on the lagged depending variable is restricted to 1. Difference-in-differences is, therefore, more restrictive than the value-added model, and in the context of the cognitive achievement producing function it implies that there is no depreciation in human capital between baseline and endline. The difference-in-differences results, shown in column (1) of Table 4, indicate that while the parameters are smaller than the results in Table 3, the differences are small and all the parameters are statistically significant.

Column (2) shows results for entropy matching. Entropy matching produces coefficients that are similar in size to those in Table 3, although for the reading score the coefficient is not significant at standard levels of confidence (p-value=0.13). To match treatment and control groups we minimize differences in the first three moments of the distributions of the same variables used in Equation (1). We also explore results using as matching variable the baseline of the dependent variable only, as well as other matching methods, based on propensity scores. These results are discussed in Annex III.

In column (3) of Table 4, we show the LASSO regression results. We use an expanded set of control variables and a LASSO regression to select the predictors. For all three outcomes of interest, the coefficients are a bit smaller than those in Table 3, but all three are statistically significant.

Overall, the results of the evaluation are robust to different specifications and the inclusion of control variables.

In Table 5, we explore treatment effect heterogeneity by gender, whether the child's caregiver can read, baseline reading skills, and grade. The interaction terms for female and the treatment

Table 4. Impact of MA on reading skills – robustness checks.

	Diff-in-diff (1)	Entropy matching (2)	LASSO (3)
PCA reading index	0.14**	0.17	0.18***
	(0.051)	(0.11)	(0.049)
	[0.00]	[0.17]	[0.00]
Oral reading fluency (cwpm)	3.16***	2.94*	2.87***
	(0.66)	(1.31)	(0.58)
	[3.65]	[9.12]	[7.40]
Reads at least one word	0.067*	0.094*	0.075**
	(0.026)	(0.042)	(0.024)
	[0.15]	[0.48]	[0.44]
N	1968	1968	1965

Notes: Regressions on the PCA reading index have five fewer observations due to item specific missing data in EGRA subskills. Outlier values, one for oral reading fluency and two for non-word decoding, both at baseline, are set to missing. The diff-in-diff regressions include the baseline value of the dependent variable, student's age and its square, gender and grade, caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, and an indicator variable for whether ChiNyanja is the language most commonly used at home, and the household asset index. Entropy matching minimizes differences between treatment and control groups of the first three moments of the distributions of the same covariates included in the diff-in-diff regressions. LASSO regression considered the following additional control variables: The five EGRA subtasks scores, district fixed effects, language most commonly used at home dummies, number of children books in ChiNyanja, four variables to account for household composition by age (0–5-year-olds, 6–15-year-olds, 16–59-year-olds, and 60-year-olds and older), an indicator for whether the house's floor is made of mud or dirt, self-reported child's school attendance in the previous week, and an indicator variable for being the first born. Standard errors clustered at the school level in parentheses. Mean value of the dependent variable for the control group at endline in squared brackets.

*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

Table 5. Treatment heterogeneity by gender, parental literacy, baseline reading skills, and grade.

	(1)	(2)	(3)	(4)
A. PCA reading index				
Treatment	0.18**	0.095	0.19***	0.098
	(0.052)	(0.066)	(0.048)	(0.061)
Interactions with treatment variable				
Female	0.025			
	(0.072)			
Caregiver can read		0.13		
		(0.071)		
Baseline value of the dependent variable			−0.0014	
			(0.061)	
3 grade student				0.19*
				(0.087)
Dep. variable mean for the control group		0.00		
B. Oral reading fluency (wpm)				
Treatment	2.67***	1.79*	2.87**	2.36*
	(0.76)	(0.87)	(0.87)	(0.93)
Interactions with treatment variable				
Female	1.32			
	(0.99)			
Caregiver can read		2.18*		
		(0.99)		
Baseline value of the dependent variable			0.11	
			(0.18)	
3 grade student				2.06
				(1.08)
Dep. variable mean for the control group		7.40		
C. Reads at least one word				
Treatment	0.095**	0.056	0.11**	0.049
	(0.028)	(0.044)	(0.032)	(0.034)
Interactions with treatment variable				
Female	0.0016			
	(0.040)			
Caregiver can read		0.056		
		(0.046)		
Baseline value of the dependent variable			−0.038	
			(0.038)	
3 grade student				0.097*
				(0.038)
Dep. variable mean for the control group		0.44		

Regressions on the PCA reading index have five fewer observations due to item specific missing data in EGRA subskills. Outlier values, one for oral reading fluency and two for non-word decoding, both at baseline, are set to missing. All regressions include the baseline value of the dependent variable, student's age and its square, gender and grade, caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, and an indicator variable for whether ChiNyanja is the language most commonly used at home, and the household asset index. Standard errors clustered at the school level in parentheses.

*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

variable are small and not significant for any outcomes. Interactions between treatment and caregiver s' literacy are positive but statistically significant for oral reading fluency only. This contrasts with Knauer et al. (2020), that find that children if illiterate parents benefit most from a parent–child book-sharing intervention in rural Kenya. We do not find significant effects when interacting treatment with baseline scores. Finally, the interaction term for third grade and the treatment dummy is positive and significant for the reading score and for whether the child can read at least one word, but only marginally significant for oral reading fluency (p-value=0.06).[17]

4.2. Reading habits and attitudes towards reading

To analyze the channels through which the program affected reading skills, we analyze the impact of the program on how often children read at home on their own or with someone else. Table 6 shows

Table 6. Impact of MA on reading habits.

	Caregiver survey		Student survey
	Number of days per week[a]	Minutes per session for min>0[b]	Number of days per week[c]
A. Read on their own	0.35**	0.59	0.24*
	(0.12)	(1.39)	(0.10)
	[2.41]	[22.5]	[1.92]
	N=1862	N=803	N=1892
B. Read with someone at home	0.56***	1.85	0.043
	(0.10)	(1.30)	(0.091)
	[2.01]	[29.9]	[1.80]
	N=1887	N=914	N=1874

Note: All models include the baseline value of the dependent variable, student's age and its square, gender and grade, caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, and an indicator variable for whether ChiNyanja is the language most commonly used at home, and the household asset index. Standard errors clustered at the school level in parentheses. Mean value of the dependent variable at endline for the control group in squared brackets.
[a]Interval regressions. Parameters can be interpreted in number of days. The reported endline mean is the mean predicted by an interval regression model with only the intercept as a covariate.
[b]OLS regressions.
[c]Negative binomial regressions.
*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

the effects of the MA program on these indicators. The estimated models follow the structure in Equation (1), and therefore in each regression, we included the baseline value of the corresponding dependent variable as a control.

Panel A shows results for the frequency in which children read on their own at home. In column (1) we use caregivers' reports. We use interval regressions to accommodate the caregiver's survey response structure. The parameter indicates that children in the treatment group read on their own at home 0.35 days per week more than children in the control group, roughly one day more every three weeks. In column (2), we show the impact of the program on the number of minutes each reading session lasted on average, restricted to cases where caregivers reported a positive number of minutes. There are no significant differences between treatment and control groups. Results using the information reported by the children are shown in column (3). We asked children how many days they had read on their own the previous week. To model these responses, we used a negative binomial regression. The results in column (3) indicate that the program increased by 0.24 days the number of days that children read at home on their own. To keep the student survey simple, we did not ask the duration of the reading sessions. While the parameters on the number of days that children read on their own at home derived from the child and caregiver surveys are different, they are both positive and significant, and the actual difference between the two coefficients is only about one standard error. These results indicate that the program had a positive impact in the number of days that students read at home on their own.

In Panel B, we show the impact of the MA program on the frequency in which children read with someone at home. Column (1) shows the estimate using as dependent variable the caregiver's responses. As before, we use an interval regression model. The parameter indicates that caregivers in the treatment group read with the children 0.56 more days per week than their counterparts in the control group; this is approximately 1 additional day every two weeks. We also asked parents about the average duration of these reading sessions. No impacts on session duration are observed according to the OLS results displayed in column (2). Finally, in column (3), we show estimates using the children' reports. The resulting coefficient is small and not significant. Therefore, our estimates indicate conflicting results between the caregiver and student reports when the outcome of interest is reading with someone at home.

It is puzzling that the effect of MA on joint reading as reported by the caregiver indicates a clear positive impact, while the effect on the same underlying construct but as reported by the children suggests no effect. One possible explanation is that while the caregiver is answering how often s/he reads with the child, the child answered how often any family member read with him or her at home, making the

two responses not strictly comparable. It is also possible that children in our sample are too young to accurately report how often they read with others at home; or that caregivers in the treatment group tend to overstate how frequently they read with the child, if they feel there is an expectation that they behave in a particular way. In this case, the challenges associated with capturing the reading frequency impedes the proper estimation of the effect of the program on joint reading at home.

Table 7 shows the impact of MA on attitudes towards reading and listening to stories using linear probability models. In column (1) the dependent variable is an indicator for whether children like listening to stories, and in column (2) for whether they like reading. Both coefficients are positive but small and not statistically significant. Similar results are found when the dependent variables are whether children like reading at home on their own, and with a relative (not shown). At endline, most children in both treatment and control groups reported that they like listening to stories and reading. For example, 91 percent of children in the treatment group said they like reading or practice reading at endline, while the figure for the control group was 90 percent. This implies that the low reading proficiency we observed is not related to lack of interest, and that the program had little room to improve reading attitudes, at least as measured by these questions.

4.3. Cost effectiveness

In this section, we estimate the program costs per student assuming the project is scaled up in the ChiNyanja-speaking areas where the pilot took place and in the rest of the country -where there are six other local languages of instruction- separately. We consider the target population to be the children in grades 2 and 3 that have access to a cell phone at home. Based on the data collected for this evaluation, we judge 60 percent a reasonable estimate of the cell phone ownership rate. We estimate that the program could reach 186,000 2nd and 3rd graders in Lusaka and the Eastern province, and 450 thousand in the rest of the country.[18]

One major cost of this program was developing the SMS stories. This alone costed almost USD 100,000, equivalent to USD 83 per treated student. However, story development is a fixed cost, and thus the associated average costs go down with the number of students treated. Moreover, this is a sunk cost for areas where ChiNyanja is the official local language (i.e. Lusaka and the Eastern province), as the stories have already been developed. Table 8 summarizes the costs of scaling up the MA program. The estimated scale-up costs per student are USD 20.1 for students in ChiNyanja-speaking areas, and USD 21.4 for students in other parts of the country.

The SMS transmission represents only about 15 percent of total costs. The two main factors driving costs are program management and outreach (of which 96 percent correspond to community mobilizers' compensation), and voice message transmission. Although the experiment design does not allow us to estimate the contribution of each program component, one program component that could be considered for removal is the voice messaging. These messages are not an obvious input for improving reading skills and were included in the program design so illiterate caregivers were more engaged in program activities; however, while about ¾ of caregivers reported they listen to the voice messages, we did not find a correlation between listening to the voice messages and caregivers' literacy.

Table 7. Impact of MA on attitudes towards reading.

	Child likes to listen to stories	Child likes to read or practice reading
Impact	0.025	0.0086
	(0.015)	(0.018)
	[0.92]	[0.90]
	N=1840	N=1860

All models include the baseline value of the dependent variable, student's age and its square, gender and grade, caregiver's literacy, household size, whether there are children's books in English or ChiNyanja at home, and an indicator variable for whether ChiNyanja is the language most commonly used at home, and the household asset index. Standard errors clustered at the school level in parentheses. Mean value of the dependent variable at endline for the control group in squared brackets.
*$p < 0.05$ **$p < .01$ ***$p < 0.001$.

Table 8. Costs per student (in USD).

Program management/outreach	$ 8.9
Notebook	$ 0.3
SMS transmission	$ 3.1
Voice message transmission	$ 7.8
Total - ChiNyanja speaking areas	$ 20.1
Story development	$ 1.3
Total - Non-ChiNyanja speaking areas	$ 21.4

(1) Cost of SMS and the voice message do not correspond to the unitary prices observed during the program, but are estimates provided by the project implementer; (2) Program management/outreach includes program coordinators (one for each of the 7 official languages in the country) and community mobilizers to contact parents to capture cell phone numbers and conduct monthly meetings with them; (3) For each of the 6 official languages for which no stories have been developed, the costs of developing stories is imputed using the corresponding costs observed during the pilot.
Source: Own calculations using data provided by Creative Associates.

5. Conclusions

This paper reports the results of a short intervention that used SMS messaging, paired with parental outreach, to improve early grade reading skills. By sending SMS stories, the program increased the access to reading materials across a population that otherwise lacks appropriate reading resources for young children. Nudging parents to have their children read, both through the very SMS and the parental meetings, was possibly another key channel to enhance children's reading skills.

We found that the program take-up was high. According to endline data, 93 percent of households in the treatment group attended at least one MA meeting, and 81 percent showed the SMS to the MA children at least once a week.

The results indicate that the program had a positive and significant impact on reading skills. Our preferred estimates of the program impacts are 19 percent of a standard deviation for the PCA reading index, 3.35 cwpm (28 percent of a standard deviation) for oral reading fluency, and 9.6 percentage points for the fraction of students able to read at least one word from a passage. Alternative estimation approaches yield very similar results. Despite these improvements, it is important to note that the absolute levels of competence among students remain low.

We found that children were spending more time reading at home on their own. Having access to reading resources facilitates home reading. It is also possible that monthly meetings had encouraged more caregiver involvement with their children's education. However, regarding reading at home with a relative, we did not find conclusive evidence that the program had a positive impact. While according to the caregiver, children were reading more at home with relatives, the effect using children's responses is negligible. Therefore, it is not clear whether caregivers spent more time reading with children or if the mechanism was solely to allow and perhaps encourage children to take time to read alone the newly provided materials.

A few important lessons for the implementation of MA and similar programs are worth highlighting. First, in general the impact seems stronger on reading skills for 3rd graders than for 2nd graders. This suggests that students need to have some pre-reading or reading skills for programs like MA to have an impact, and that, if replicated, efforts should probably focus on those students. Second, the relevance of the voice recording should be weighed against its cost. The voice recording was included as a way for illiterate parents to engage in the program, however, we found that its use was not correlated with caregivers' literacy; being twice as expensive as the SMS, it is reasonable to see this program component as a candidate for removal. Third, while the SMS stories are fundamental aspect of the program, caregiver's engagement encouraged through the meetings should not be underestimated. Our findings suggest that both components of the program were important in improving students' reading performance. Finally, MA requires mobile phone access and therefore it cannot reach the poorest children who do not have it. Scaling-up the program could potentially

increase education inequalities between the very poor and better-off children with phones, or between geographic areas with different connection coverage. However, as mobile phones and connection plans are becoming cheaper and connectivity is improving across the country, this potential issue will become less important.

Notes

1. In Sub-Saharan Africa, for example, there were 74 cell phone subscriptions per 100 people in 2017 (World Bank 2018a).
2. *Makhalidwe Athu* can be translated as "our way of life", "our way of staying" or "our culture".
3. Many studies document the impact of using SMS messaging as reminder vehicle in the context of health-related interventions. In developing countries this includes, for example, programs to improve adherence to antiretroviral therapy (Kanters et al. 2017), diabetes self-care (Abaza and Marschollek 2017), and health workers' adherence to malaria treatment guidelines (Zurovac et al. 2011) among many others types of programs.
4. The seven official languages of instruction are: ChiNyanja, Bemba, Lozi, Tonga, Kaonde, Luvale, and Lunda.
5. A total of 57 schools were dropped from the original list of 350 schools that these two districts have, due to missing data (25), poor cell phone coverage or low accessibility (24), or because they were too small (8).
6. The random selection of 80 schools used distance to District Education Board Secretary and school size for stratification. These variables were also used to match schools before randomization.
7. Because the sample was constructed at the end of the school year (October 2015), we interviewed students in grades 1 and 2 so most of them would transition to grades 2 and 3, respectively, in 2016, when the program was going to treat them.
8. In a handful of very large schools not all students were screened but a random sample.
9. Some respondents were parents of more than one target child, which is why we have more student-level caregiver data than caregiver respondents.
10. The respondent was a parent 80 percent of the cases.
11. To estimate whether there are significant differences between treatment and control for ordered categorical variables (caregivers' responses on how often they read with the child and how often the child reads alone) we estimated ordered probit regressions where the outcome of interest is the ordered categorical variable and the covariate is the treatment dummy.
12. Kerwin and Thornton (2018) follow a similar aggregation of the EGRA scores.
13. Even if we restrict the sample to children starting 3rd grade at baseline (roughly half the sample), the average ORF score is 6.8 cwpm (pooling both study groups), which is also way below the minimum threshold of 25 cwpm expected for this grade.
14. This algorithm consists of the following steps: (i) Use Lasso to select predictors of the treatment variable; (ii) Use Lasso to select predictors of the outcome variable, and (iii) Regress the outcome of interest on the treatment variable and the covariates selected in steps (i) and (ii).
15. Descriptive statistics for oral reading fluency and other outcomes of interest at endline for the control group, as well as histograms for oral reading fluency at endline for the control group, can be consulted in Annex II.
16. Results for the other four EGRA subtasks can be consulted in Annex II.
17. Heterogeneity by grade for the PCA reading index is in part an artifact of standardizing the index using the mean and standard deviation of the pooled 2nd and 3rd graders in the control sample. In Annex II we show impact evaluation results splitting the sample by grade, using as dependent variables both the reading score standardized with the mean and standard deviation of the control group pooling grades 2 and 3, and the reading score standardized using grade-specific mean and standard deviations of the control group. The results indicate that even if we use reading scores standardized at the grade level, the difference between grades is still large and statistically significant.
18. Using census data for the total and by-region population (Central Statistical Office 2012), UNESCO data for the number of 2nd and 3rd graders (UNESCO Institute for Statistics 2019), and our own survey data for the cell phone ownership rate, we estimate that by 2016 the number of 2nd and 3rd graders in the country was 1 million, and of those 636 thousand had access to a cell phone at home. The Eastern Province and Lusaka account for approximately 29 percent of the country's population.

Acknowledgements

The authors' views expressed here do not necessarily reflect the views of USAID. Although USAID staff provided comments during the different stages of the research, the authors were the sole responsible for the study design, data collection, analysis and interpretation of data, and the writing of the report. Similarly, the decision to submit the article for publication was the authors' only. The authors would like to thank participants at the Association for Public Policy

Analysis and Management conference and the University of Chicago Demography Workshop for their valuable comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the United States Agency for International Development (USAID) Contract No. GS-10F-0033M/AID-OAA-M-13-00010.

References

Abaza, Haitham, and Michael Marschollek. 2017. "SMS Education for the Promotion of Diabetes Self-Management in Low & Middle Income Countries: A Pilot Randomized Controlled Trial in Egypt." *BMC Public Health* 17 (1): 962. doi:10.1186/s12889-017-4973-5

Abdulkadiroglu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2011. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics* 126 (2): 699–748. doi:10.1093/qje/qjr017

Aker, Jenny C., Christopher Ksoll, and Travis J. Lybbert. 2012. "Can Mobile Phones Improve Learning? Evidence from a Field Experiment in Niger." *American Economic Journal: Applied Economics* 4 (4): 94–120. doi:10.1257/app.4.4.94.

Belloni, A., V. Chernozhukov, and C. B. Hansen. 2014. "Inference on Treatment Effects After Selection among High-Dimensional Controls." *Review of Economic Studies* 81: 608–650. doi:10.1093/restud/rdt044

Beuermann, Diether W., Julian Cristia, Santiago Cueto, Ofer Malamud, and Yyannu Cruz-Aguayo. 2015. "One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru." *American Economic Journal: Applied Economics* 7 (2): 53–80. doi:10.1257/app.20130267

Brombacher, Aarnout, Jennae Bulat, Simon King, Emily Kochetkova, and Lee Nordstrum. 2015. *National Assessment Survey of Learning Achievement at Grade 2. Results for Early Grade Reading and Mathematics in Zambia*. Research Triangle Park, NC: RTI International.

Bulman, George, and Robert W. Fairlie. 2016. "Technology and Education: Computers, Software, and the Internet." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 5: 239–280. Amsterdam: North-Holland.

Central Statistical Office. 2012. "2010 Census of Population and Housing, Volume 11, National Descriptive Tables." Lusaka: Central Statistical Office, Nationalist Road P.O. Box 31908.

Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín. 2017. "Technology and Child Development: Evidence from the One Laptop per Child Program." *American Economic Journal: Applied Economics* 9 (3): 295–320. doi:10.1257/app.20150385

Deming, David J. 2014. "Using School Choice Lotteries to Test Measures of School Effectiveness." *American Economic Review: Papers & Proceedings* 104 (5): 406–411. doi:10.1257/aer.104.5.406

Deutsch, Jonah. 2012. "Using School Lotteries to Evaluate the Value-Added Model." Unpublished manuscript. University of Chicago.

Graham, J., and S. Kelly. 2018. "How Effective are Early Grade Reading Interventions?" A Review of the Evidence. Policy Research Working Papers. https://doi.org/10.1596/1813-9450-8292.

Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20: 25–46. doi:10.1093/pan/mpr025

Hanemann, Ulrike, and Cassandra Scarpino, eds. 2016. *Harnessing the Potential of ICTs: Literacy and Numeracy Programmes Using Radio, TV, Mobile Phones, Tablets and Computers*. Hamburg, Germany: UNESCO Institute for Lifelong Learning.

Kaleebu, Nasiib, Alison Gee, Amanda H.A. Watson, Richard Jones, and Marshall Jauk. 2017. "SMS Story: A Case Study of a Controlled Trial in Papua New Guinea." In *Mobile Learning in Higher Education in the Asia-Pacific Region*, edited by Angela Murphy, Helen Farley, Laurel Evelyn Dyson, and Hazel Jones, 623–645. Singapore: Springer Nature Singapore.

Kanters, Steve, Jay J. H. Park, Keith Chan, Maria Eugenia Socias, Nathan Ford, Jamie I. Forrest, Kristian Thorlund, Jean B. Nachega, and Edward J. Mills. 2017. "Interventions to Promote Adherence to Antiretroviral Therapy: a Systematic Review and Network Meta-Analysis." *The Lancet. Hiv* 4: e31–e40. doi:10.1016/S2352-3018(16)30206-5

Kerwin, Jason T., and Rebecca L. Thornton. 2018. "Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures." Unpublished manuscript.

Knauer, Heather A., Pamela Jakiela, Owen Ozier, Frances Aboud, and Lia C.H. Fernald. 2020. "Enhancing Young Children's Language Acquisition Through Parent–Child Book-Sharing: A Randomized Trial in Rural Kenya." *Early Childhood Research Quarterly* 50 (Part 1): 179–190. doi:10.1016/j.ecresq.2019.01.002.

Malamud, Ofer, and Cristian Pop-Eleches. 2011. "Home Computer Use and the Development of Human Capital." *Quarterly Journal of Economics* 126 (2): 987–1027. doi:10.1093/qje/qjr008

Mayer, Susan E., Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2018. "Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together Intervention." *Journal of Human Resources. Published Ahead of Print*. doi:10.3368/jhr.54.4.0617.8835R.

McEwan, Patrick J. 2015. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments." *Review of Educational Research* 85 (3): 353–394. doi:10.3102/0034654314553127

Ministry of Education, Science, Vocational Training and Early Education. 2014. "Reading Performance Level Descriptors for Grades 1–4." Directorate of Standards & Curriculum, in Collaboration with USAID/Zambia Read to Succeed Project.

Murray, N., L. Campuzano, C. Fernandez, J. Lugo-Gil, S. Luizzi, E. Bagby, and I. Padilla. 2021. "The Future of Literacy in Latin America: Evidence from LAC Reads" https://www.mathematica.org/events/the-future-of-literacy-in-latin-america-how-evidence-from-lac-reads-can-help-close-early-reading.

Pouezevara, Sarah, and Simon King. 2014. "MobiLiteracy-Uganda Program Phase 1: Endline Report." Research Triangle Park, NC: RTI International.

Reid, Danielle, and Christopher Pruijsen. 2015. "Increasing Learning Outcomes in Developing Countries by Engaging Students Out of Classroom Using SMS and Voice Mobile Technology." In *Handbook of Mobile Teaching and Learning*, edited by Yu (Aimee) Zhang, 421–436. Berlin: Springer.

RTI International. 2015. *Early Grade Reading Assessment (EGRA) Toolkit, Second Edition*. Washington, DC: United States Agency for International Development.

Singh, A. 2015. "Private School Effects in Urban and Rural India: Panel Estimates at Primary and Secondary School Ages." *Journal of Development Economics* 113: 16–32. doi:10.1016/j.jdeveco.2014.10.004

UNESCO Institute for Statistics. 2019. Data From: Enrolment in Primary Education (Grades 2 and 3).

World Bank. 2018a. World Bank Open Data. World Bank, Washington, DC. © World Bank. Accessed July 26, 2020. https://data.worldbank.org/indicator/IT.CEL.SETS.P2.

World Bank. 2018b. The Human Capital Project. World Bank, Washington, DC. © World Bank.

York, Benjamin N., Susanna Loeb, and Christopher Doss. 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. doi:10.3368/jhr.54.3.0517-8756R

Zurovac, Dejan, Raymond K. Sudoi, Willis S. Akhwale, Moses Ndiritu, Davidson H. Hamer, Alexander K. Rowe, and Robert W. Snow. 2011. "The Effect of Mobile Phone Text-Message Reminders on Kenyan Health Workers' Adherence to Malaria Treatment Guidelines: A Cluster Randomized Trial." *Lancet* 378: 795–803. doi:10.1016/S0140-6736(11)60783-6