

# Nudges to improve learning and gender parity: Supporting parent engagement and Ghana's educational response to Covid-19 using mobile phones

## Pre-analysis Plan

*While recent evidence from Brazil and Ivory Coast suggests that SMS messages to nudge parents' engagement in their children's education can improve on educational outcomes, the Covid-19 pandemic raises additional concerns. In particular, learning deficits and school dropouts are likely to increase following school shutdowns, especially among vulnerable populations such as older girls who need to work to support their families or due to early marriage, childbearing and adolescent pregnancy. A further knowledge gap relates to the optimal period of exposure of nudges, which is critical to scale-up. This study investigates whether sending nudges to parents can improve parental engagement in child education as well as broader development across child age groups and gender in rural Ghana by randomly assigning whether parents receive two different versions of nudges, with one version including content promoting girls' education and addressing some common stereotypes around gender, and whether the duration of these different modalities vary between three and six months.*

## I. Introduction

Ghana's school reopened in January 2021 after a 10-month period of a remote-learning program aimed at reaching students through Radio, Television, and Internet to ensure continuity of learning throughout the crisis. For children to be successful as they re-enter educational activities, parent involvement is essential; yet how to best communicate with parents, and whether parents with low education levels can support remote learning, is unclear.

Ghana's Education Strategic Plan (ESP) and the Ghana Accountability Learning Outcomes Project (GALOP) recognizes parents/guardians as key stakeholders in children's education and aims to develop a communications strategy to deliver key messages to teachers, parents, and students. While parents are required by law to enroll their children in school, the level of involvement in children's education is generally low, particularly in the poorer regions in northern Ghana.

Further, the health and economic crises resulting from COVID-19 affected households' economic situations, especially the most vulnerable ones. It is possible that many children, particularly older girls, will not return to school after the crisis because of the need to work to support their families or due to early marriage and childbearing, and adolescent pregnancy.

Ghana's Human Capital Index is 0.44, meaning that a child born today can only be expected to reach 44% of his/her potential. With schools closed, the situation may become more dire. The need to

find low-cost, gender-sensitive solutions to minimize disruptions to learning and schooling is urgent, especially in most-disadvantaged northern regions.

Against a low learning base, the effectiveness of remote-learning, and its role in widening or mitigating inequalities, especially of gender, remain open. For many children, the ability to learn during the crisis and beyond will critically depend on their parents' engagement. However, parents often face informational barriers to support learning effectively. Further, parent engagement may vary by child gender, due to greater opportunity costs of schooling for girls (e.g. larger involvement of girls in household or care-work), lower perceived returns to girls' education, and widespread gender bias in social norms and aspirations. Providing timely, actionable information to poor and low-educated parents, including via text-messages as a low-cost intervention, can attenuate these barriers and improve parental engagement across child-age groups and gender (Bergman, 2019). If such interventions work during and after the pandemic, where stressors are greater than under non-emergency circumstances, and in a low-resource setting, is unknown.

This pre-analysis plan summarizes the design of a field experiment designed to test **the following primary hypotheses:**

1. Do behavioral nudges to parents in the form of SMS-messages increase parent engagement in their children's education?

- Hypothesis: SMS nudges increase parental engagement in their child's education and school life.

2. Do messages change parental beliefs about returns to education and educational expectations and aspirations for each child of target age?

Hypothesis: SMS nudges increase parents' beliefs about the returns of education as well as their support for and investment in their child's education and aspirations for the future.

3. Do messages improve children's enrollment and attendance in school post-schools reopening?

- Hypothesis: SMS nudges increase (or at least dampen the negative effects of school shutdowns on child school enrollment and attendance.

- Assuming parental educational investments and school attendance increase, it is not clear that children's learning outcomes will improve. Given low educational quality, attending school does not necessarily translate into improved learning. This question assesses if nudges impact other dimensions of children's development that have potential prospects for longer-term well-being.

4. Are these impacts more equitable across girls and boys if the messages additionally focus on gender-parity in education and in behaviors/attitudes towards girls?

- Hypothesis: Focusing on gender-parity in education and in behaviors/attitudes towards girls supports the achievement of more gender-equitable outcomes.
5. Do these impacts differ for younger (5-9 years), versus older (10-17 years) children?
- Hypothesis: Based on their greater involvement in child labor (within or outside the household), we hypothesize that the intervention may have a stronger effect on older children by increasing the time spent in educational activities vis-à-vis time spent in labor. However, given the on-going crisis, it may be the case that older children may increase even more their work to support the family as the educational opportunity costs may have raised substantially (e.g. caring for younger siblings, or greater involvement in farm or business), so that the intervention may have a larger effect on younger children as compared to the older ones.
6. Are these impacts larger and do they persist for longer if delivered for a longer duration (6 versus 3 months)?
- Hypothesis: Increasing the duration of delivery of SMS nudges increases the impacts over time. If parents continue to be nudged, impacts in the next school year will increase.

## **II. Intervention**

The Parental Nudges Project (PNP) is a household-level intervention designed to improve school-aged children’s outcomes by engaging parents in their children’s education. The intervention involves two text-messages (SMS) per week sent to parents / primary caregivers in simple English with behavioral “nudges” around engaging with children’s learning generally and during remote learning across grades and ages. Messages include suggestions of simple activities that promote nurturing child social-emotional development and education. No curricular knowledge is required. During school closures, suggestions related to remote learning will be included.

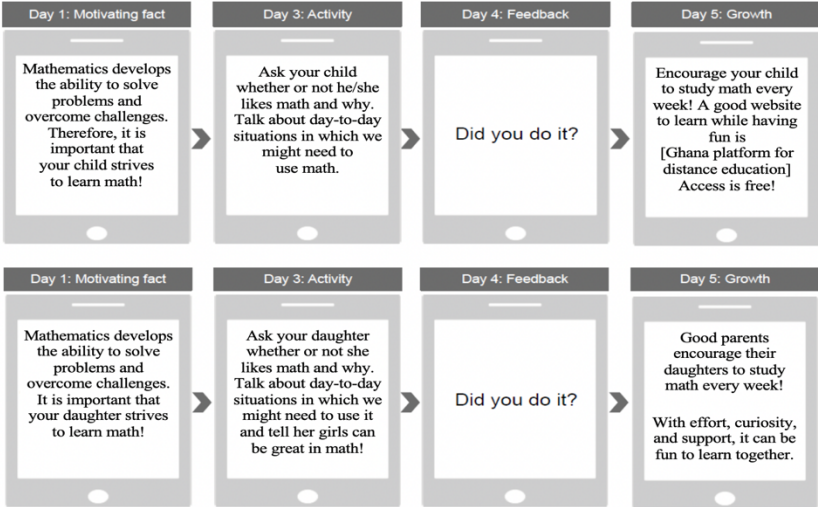
As school-age girls (ages 5-17 years) in primary schools are likely to be disproportionately affected by the crisis, a randomized subset of parents will receive messages promoting gender-equitable outcomes. Nudges will encompass reminders, encouragement and activities addressing information gaps, biased beliefs, and norms behind gender inequalities in education and broader development.

We will also vary the duration of exposure to the interventions across treatment arms. While the general parental engagement intervention has been implemented in several countries, the effectiveness of the gender component and the differing duration of exposure to messages in inducing belief and behavior change has never been tested.

The intervention is nimble, and message contents can be changed rapidly. We have adapted content as schools re-opened (January 2021) and align them to government and World Bank remote-learning and back-to-school campaigns.

All nudges will be sent from a short code number (a 5-digit number), and include the EDU+/IPA tag, to clearly identify the messages under this initiative. Edu+ (powered by EdTech Movva) shares weekly suggestions of activities for parents to do with their children – none of them linked to curricular activities; rather, those try to bring parents closer to their children’s school life by having them ask about school, discuss future plans, and share how they dealt with similar conflicts back in the day. Nudges are structured around sequences in a format inspired by READY4K!, an eight-month-long text-messaging intervention for parents of preschoolers that targets the behavioral barriers to engaged parenting (York et al., 2017). The figure below showcases two examples of the SMS sequence sent to parents assigned to the nudge program: the first sequence is not gender specific, while the second is specifically targeted at tackling gender inequalities.

Figure 1: Two examples of SMS sequence sent to parents



### III. Experimental Design

The unit of randomization is households. To assign households to treatment and control, we will employ a household-level randomized controlled trial design. Households in the control group will not receive any messages during the study period. There will be no stratification as part of the randomization. Eligible and consented households identified through the Enrollment Call will be randomly assigned to receive one of the four treatment SMS text message groups or to receive no SMS text message. The randomization protocol, which will be implemented through a STATA do-file, will seek to achieve a

1:1:1:1 ratio across the five experimental groups. The use of a STATA do-file is to ensure that the randomization is reproducible.

The experimental groups are:

1. Treatment group 1. Behavioral intervention: Nudges to parents supporting involvement with children’s learning, their child’s social-emotional development, academic aspirations, and engagement in remote learning activities during the school closures and into the summer (3 months);
2. Treatment group 2. A “gender-equality boost” arm, in which some of the nudges include content promoting girls’ education and addressing some common stereotypes around gender roles during the school closures and into the summer (3 months);
3. Treatment group 3. Treatment 1 implemented for 6 months into the first term of the next academic year;
4. Treatment group 4. Treatment 2 implemented for 6 months, into the first term of the next academic year;
5. Control group. No intervention / no messages.

The following surveys will be conducted (dates may change based on the evolution of the pandemic):

- Enrollment and Baseline Phone Survey – December 2020
- Implementation Phone Survey I – February 2021
- Child In-person Assessment I (N=5000) – April-May 2021
- Parent In-person Assessment I (N=2,500)– April 2021-May 2021
- Child In-person Assessment II (N=5000) – September-October 2021
- Parent In-person Assessment II (N=2,500)– September-October 2021

All assessments will be translated into the various local languages of the study regions and simultaneous administrative data on children’s attendance will also be collected in tandem with the final assessments.

The intervention will be evaluated through a household-randomized controlled trial with a goal of 2,500 households (500 in each treatment arm) in the poorest regions in Ghana (Northern, Upper East and Upper West Regions) – the exact number is still to be determined. Randomizing at the household level ensures that we take into account within-household spillovers. Within each household, beyond the parent, we will sample two children to investigate within-household spillovers and age- and gender-heterogeneity in key study outcomes: one child in the lower basic education (including early-childhood/early primary grades) (5-9 years) and one child in the upper basic education and adolescence

age range (10-17 years). If there are more than two children of those target ages within the household, we will randomly select one based on a household roster.

Each household will be assigned to one of the five treatment/control groups, as shown in table 1 below:

Table 1: Randomization strategy

	<b>Yes (3 months)</b>	<b>Yes (6 months)</b>	<b>No</b>
<b>Behavioral nudges</b>	500 households 500 caregivers 1,000 school-age children	500 households 500 caregivers 1,000 school-age children	500 households 500 caregivers 1,000 school-age children
<b>Gender-equality boost</b>	500 households 500 caregivers 1,000 school-age children	500 households 500 caregivers 1,000 school-age children	

Our sample is drawn from two previously completed studies. First, an impact evaluation of the Communications for Development (C4D) study<sup>1</sup> (2012-2016), launched by the Ghana Health Service with funding from UNICEF in 12 districts of the three poorest regions of Ghana (these were, at the time, Northern, Upper East, and Upper West, and later they have split in further regions). The sample included mothers with a child aged 0-5 years recruited in 2012. The program relied on voice messages directly delivered to female respondents through their cell phones, and the sample has high rates of mobile phone ownership (83%). In 2016, there were over 4300 families with children under age 7 years in this sample, indicating that we will have sufficient sample size for our target 2,500 families, even in a very negative scenario in which 40% of them have changed phone number and we cannot trace them in other ways<sup>2</sup>. Second, we will be employing households from a subsample of the Ghana Panel Survey<sup>3</sup> (specifically, the Graduating Out of Poverty sub study) from the same regions to obtain our desired sample size.

In summary, the sample size for this study takes into account the required sample for statistical inference, possible attrition rates, and piloting of the intervention. The sample comprises of 2,500 households across the five experimental groups, 2,500 caregivers and 5,000 school-age children (2 from each household) from the selected households. Current funding permits us to conduct phone-surveys with 3,550 children, and in-person assessments for 1,000 children. With such sample sizes, we will be able to detect the following treatment effects:

<sup>1</sup> <https://www.poverty-action.org/study/communication-development-ghana>

<sup>2</sup> Note: The figure of 40% untraceability rates is based on evidence of attrition in Sierra Leone and Liberia during the Ebola crisis.

<sup>3</sup> <https://www.src.isr.umich.edu/international/ghana-socioeconomic-panel-study/>

PARENT-LEVEL OUTCOMES		
	<b>Sample size</b>	
<b>Treatment 1</b>	500 households	
<b>Treatment 2</b>	500 households	
<b>Treatment 3</b>	500 households	
<b>Treatment 4</b>	500 households	
<b>Comparison</b>	500 households	
<b>Assumed attrition</b>	10%	
	<p><b>Power: 0.80</b></p> <p>For the first follow-up (midline), we will pool treatments 1 and 3, and treatments 2 and 4 to double our sample size.</p> <p>For the second follow-up (endline), where we compare differences in duration of exposure, we will analyze these arms separately. We provide estimates of MDES for both.</p> <p>We assume an R-squared of 0.1, a conservative estimate given that we have three waves of data over five years.</p>	<p><b>Power: 0.90</b></p> <p>For the first follow-up (midline), we will pool treatments 1 and 3, and treatments 2 and 4 to double our sample size.</p> <p>For the second follow-up (endline), where we compare differences in duration of exposure, we will analyze these arms separately. We provide estimates of MDES for both.</p> <p>We assume an R-squared of 0.1, a conservative estimate given that we have three waves of data over five years.</p>
<b>Number of clusters</b>	<p><u>Midline</u>: 1350 (pooling treatments 1 and 3, and treatments 2 and 4, relative to the control).</p> <p><u>Endline</u>: 900 (for a 2-way comparison between treatment and control)</p> <p>(after 10% attrition accounted for)</p>	<p><u>Midline</u>: 1350 (pooling treatments 1 and 3, and treatments 2 and 4, relative to the control).</p> <p><u>Endline</u>: 900 (for a 2-way comparison between treatment and control)</p> <p>(after 10% attrition accounted for)</p>
<b>Size of cluster</b>	1 parent	1 parent

<b>Significance level (alpha)</b>	0.05	0.05
<b>Sample size</b>	<u>Midline</u> : 1350 <u>Endline</u> : 900 for parents	<u>Midline</u> : 1350 <u>Endline</u> : 900 for parents
<b>Minimum detectable effect</b>	<u>Midline</u> : 0.14 SD <u>Endline</u> : 0.18 SD	<u>Midline</u> : 0.17 SD <u>Endline</u> : 0.21 SD
<b>CHILD-LEVEL OUTCOMES</b>		
	<b>Sample size</b>	
<b>Treatment 1</b>	500 households, 1,000 children	
<b>Treatment 2</b>	500 households, 1,000 children	
<b>Treatment 3</b>	500 households, 1,000 children	
<b>Treatment 4</b>	500 households, 1,000 children	
<b>Comparison</b>	500 households, 1,000 children	
<b>Assumed attrition</b>	10%	
	<b>Power: 0.80</b>  For the midline, we will pool treatments 1 and 3, and treatments 2 and 4 to double our sample size.  We assume an R-squared of 0.1, a conservative estimate given that we have three waves of data over five years. For children nested in households, we assume an ICC of 0.10 given the wide age range.	<b>Power 0.90</b>  For the midline, we will pool treatments 1 and 3, and treatments 2 and 4 to double our sample size.  We assume an R-squared of 0.1, a conservative estimate given that we have three waves of data over five years. For children nested in households, we assume an ICC of 0.10 given the wide age range.
<b>Number of clusters</b>	<u>Midline</u> : 1350 households (assuming 10% attrition) <u>Endline</u> : 990 (assuming 10% household attrition)	<u>Midline</u> : 1350 households (assuming 10% attrition) <u>Endline</u> : 990 (assuming 10% household attrition)
<b>Size of cluster</b>	2 children / household	
<b>Significance level (alpha)</b>	<b>0.05</b>	<b>0.05</b>



<b>Sample size</b>	<u>Midline</u> : 1,980 (assuming 10% household attrition)	<u>Midline</u> : 1,980 (assuming 10% household attrition)
	<u>Endline</u> : 2,700 (assuming 10% household attrition)	<u>Endline</u> : 2,700 (assuming 10% household attrition)
<b>Minimum detectable effect</b>	<u>Midline</u> : 0.048 SD	<u>Midline</u> : 0.054
	<u>Endline</u> : 0.056 SD	<u>Endline</u> : 0.065 SD

#### IV. Measures

We will collect measures through parent / guardian surveys, child direct assessments, and school administrative records. Specifically, we will collect the following:

1. **Parent’s engagement in child schooling and learning** (both with regards to remote-learning activities and general learning when/if schools reopen)

Parent-report: We will ask through three interviews a battery of questions about parent’s engagement on child education (e.g. reading to child, talking with the child, playing, etc.), and with government remote-learning activities. We will build a summary measure of parent engagement in child education by counting the activities in which the parent report s/he is involved, similarly to Banerji et al. (2017).

2. **Parent’s educational aspirations, expectations and estimated returns on education for each child**

We ask parents to report on their educational aspirations and expectations for two focal children in the household: ‘What is the highest level of education that you WISH [child] to achieve?’, and ‘What is the highest level of education that you EXPECT [child] to achieve?’. Based on actual data we will gather, we will decide a meaningful cutoff point, though we are inclined to follow previous literature (e.g. Favara 2017) that identified high aspirations as completing tertiary education. For returns to education measures, the parent is asked to estimate their child’s employability and how much (in Ghanaian cedis) they expect their child would make given a JHS versus SHS versus University degree. This measure was based on previous literature (e.g. Attanasio and Kauffman, 2014) and then field tested in our sample.

3. **Gender norms**

This will be assessed three times, once during the initial enrollment call and twice in the midline and endline, in-person surveys. Gender norms will be measured through the “Gender norms and attitudes

scale” (Waszak et al, 2000), which measure egalitarian beliefs about male and female gender norms. Specifically, the scale assesses agreement/disagreement with a number of statements related to the promotion of equity for girls and women and the maintaining the rights and privileges of men (14 items/2 subscales).

#### 4. **Child schooling outcomes** (enrollment and attendance)

This will be assessed through three sources: parent-report, child-report, and school administrative records. At baseline, parents report through a phone survey whether their child was enrolled in school in the previous academic year (2019-2020). In the midline and endline surveys, we will ask both parents and children about their enrollment in school during their in-person assessments. Further, we will ask children and parents’ about their/their child’s attendance as well as collect school administrative records about each child’s enrollment and attendance.

#### 6. Self-efficacy

Parental self-efficacy is measured using one scale from Bandura’s Parental Self-Efficacy Scale (Bandura et al., 2001), specifically the items related to parental self-efficacy regarding their children’s schooling and learning. Self-efficacy is assessed during both the midline and endline parent surveys.

#### 7. Parent mental health

Parent mental health is through in-person surveys with parents using the Kessler Psychological Distress Scale (Kessler, R.C., Andrews, G., Colpe, Et al, 2002), a 10-item questionnaire used globally to measure general psychological distress based on questions about anxiety and depressive symptoms. This scale will only be used as endline.

#### 8. Discipline practices

The discipline practices used by parents is assessed using the UNICEF MICS scale (MICS, UNICEF, 2014). Some of the nudges focus on how corporal punishment can harm a parent’s relationship with their child, so this construct will be assessed at both midline and endline through in-person surveys.

#### 9. Emotional supportiveness

Parent emotional supportiveness is assessed using a scale taken from the ECLS-K parent survey (U.S. Department of Education, 2015) that ask parents a variety of hypothetical questions specific to supporting their children’s emotional needs. This data will be collected via in person parent surveys at both midline and endline.

## 10. Children's internalizing and externalizing of symptoms

To measure children's internalizing and externalizing outcomes, we use the Strengths and Difficulties Questionnaire parent-report scale (Goodman, R., 1997) at midline and endline through in-person surveys. The SDQ is a 25-item survey that assesses both positive and negative aspects of psychological wellbeing and has been tested and used in a variety of contexts, including Ghana.

At the **child level** we will measure enrollment and attendance, learning (literacy and numeracy), and developmental outcomes (social-emotional development, motivation for school, educational aspirations, self-esteem, time use, and perceived returns to education for older group) through direct assessments with children. Further, we will collect school administrative records in the fall of 2021 to assess additional data on school enrollment and attendance.

We will assess very basic literacy and numeracy knowledge. There will be different scales for children in the different age groups, which we will devise by selecting items for administration through in-person surveys from the assessments used in the PI's Quality Preschool for Ghana Study (funded by SIEF) and an impact evaluation on school feeding that was also conducted in the north of Ghana (Aurino et al., 2018), and consider questions from Ghana's national exam. We will also draw on other item data banks such as Young Lives and EGRA/EGMA. In addition, we will assess children's educational aspirations (as measured by using an adaptation of the Quality Preschool for Ghana survey module focusing on children, which asks the child which level of formal education she would like to achieve if she does not have constraints to her schooling), and current school enrollment and attendance status (the latter to confirm data we collect in the household survey). Further, we will ask older children (ages 10-17) about their time use (for younger children, we will ask their parents) by again adapting the Young Lives survey questions about the time the child has spent in a typical day in educational activities, house- and care-work, work out of the household, leisure, etc. We hypothesize that time use is an important and potentially highly gendered pathway for intervention impact.

For the younger children, we will draw from the International Development and Early Learning Assessment (IDELA) to measure literacy, numeracy, and social-emotional development through an adaptation of the International Socio Emotional Learning Assessment (ISELA). We will also use items from the EGRA/EGMA and other existing scales measuring social-emotional outcomes such as the Rosenberg Self-esteem Scale. For the older children, we will draw on data collected from a study of Aurino's in the northern regions to assess literacy and numeracy plus the Young Lives data. Further, we will measure child time use, educational aspirations, and psycho-emotional development.

## V. Outcomes

We will document the effects of the treatments on the following outcomes:

### Primary outcomes

1. Parent engagement in education
2. Parent expectations on returns on education
3. Parent expectations and aspirations for their child’s schooling
4. Parent beliefs about gender norms
5. Children’s school enrollment and attendance

### Secondary outcomes

Secondary outcomes are divided into those we hypothesize will be mediators to the primary outcomes and those that we consider distal outcomes to the treatment.

Mediators	Distal
<ol style="list-style-type: none"> <li>1. Child time use (child-report)</li> <li>2. Parental self-efficacy (parent-report)</li> <li>3. Emotional Supportiveness in the home (parent-report)</li> </ol>	<ol style="list-style-type: none"> <li>1. Child literacy and numeracy outcomes (direct child assessments)</li> <li>2. Child social and emotional skills (direct child assessments and parent report)</li> <li>3. Parental mental health (parent report)</li> </ol>

We will be able to test for balance across treatment arms on household (e.g. age of the caregiver, number of family members, physical aspect of the house as reported by the respondent, access to drinking water and other necessary resources, etc) and child development characteristics (e.g. physical/motor development, cognitive and language, social-emotional, etc) by conducting mean comparison tests. Since information on such outcomes was already collected in the context of the Communications for Development study in 2016 and for the Graduating Out of Poverty study in 2019, and we will conduct a baseline survey during the enrollment call, we will be able to test the balance across groups even before the start of this intervention.

## VI. Empirical analysis

The randomized design allows for the identification of causal effects of the interventions on parents and children by comparing mean outcomes between the randomized treatment arms. The analysis will follow an intention-to-treat approach, using econometric analysis for all the relevant outcomes of the intervention.

**For each parent outcome** described above (Section V), we will estimate the following ordinary least squares regressions indexed by parent  $p$  from household  $h$  and survey  $s$ :

$$Y_{p,h,s} = \beta_0 + \beta_1 NShort_h + \beta_2 GShort_h + \beta_3 NLong_h + \beta_4 GLong_h + \beta_5 Y_{p,h,0} + \beta_6 X_{h,0} + \beta_7 G_h + \theta_r + \varepsilon_{h,s}$$

Where:

- $Y_{p,h,s}$  is the outcome variable for parent  $p$  in household  $h$  and survey round  $s$ ;
  - $NShort$ ,  $GShort$ ,  $NLong$ , and  $GLong$  are indicator variables assuming the value of 1 if the household has been randomly assigned to any of the treatment arms (Arm1: Nudges Short duration; Arm2: Gender boost short duration; Arm3: Nudges long duration; and Arm4: Gender boost long duration). We note that for the first parental assessment (the one conducted at the end of the three-month implementation of treatments 1 and 2), we will pool samples from arms 1 and 3, and 2 and 4 to estimate effects. For the endline,  $T_p$  will include four treatment dummies to treatment arms 1, 2, 3 and 4 separately.
  - $Y_{p,h,0}$  is the baseline outcome variable for parent  $p$  in household  $h$  (when available)
  - $X_{h,0}$  is a vector of parent and household controls should there is lack of balance in the randomization
- $G_h$  is an indicator variable assuming the value of 1 for households belonging to the GUP sample, 0 otherwise<sup>4</sup>;
- $\theta_r$  are region fixed effects

$\varepsilon_{h,s}$  is individual error term

We are interested in testing  $\beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$ .

**For each child outcome** described above (Section V), we will estimate the following ordinary least squares regressions indexed by child  $c$ , living in household  $h$  and survey  $s$ :

$$Y_{c,h,s} = \beta_0 + \beta_1 NShort_h + \beta_2 GShort_h + \beta_3 NLong_h + \beta_4 GLong_h + \beta_6 X_{h,0} + \beta_7 G_h + \theta_r + \varepsilon_{c,h,s}$$

---

<sup>4</sup> The C4D sample was part of a randomized intervention trial. The treatment had no impacts. As a robustness check, we will also add an indicator for treatment status in the C4D sample.

Where:

- $Y_{c,h,s}$  is the outcome variable for child  $c$ , living in household  $h$  and survey  $s$
- NShort, GShort, NLong, and GLong are indicator variables assuming the value of 1 if the household has been randomly assigned to any of the treatment arms (Arm1: Nudges Short duration; Arm2: Gender boost short duration; Arm3: Nudges long duration; and Arm4: Gender boost long duration). As for the parents' outcomes, for the first child assessment (the one conducted at the end of the three-month implementation of treatments 1 and 2), we will pool samples from arms 1 and 3, and 2 and 4 to estimate effects. For the endline,  $T_p$  will include four treatment dummies to treatment arms 1, 2, 3 and 4 separately.
- $X_{h,0}$  is a vector of parent and household controls should there is lack of balance in the randomization
- $G_h$  is an indicator variable assuming the value of 1 for households belonging to the GUP sample, 0 otherwise<sup>5</sup>;
- $\theta_r$  are region fixed effects
- $\varepsilon_{c,h,s}$  is individual error term, clustered at the household level.

We are interested in testing  $\beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$ .

### Heterogeneity

We will assess differences in treatment effects by child gender and age group. We will also evaluate heterogeneity in treatment effects by main parents' educational level, baseline household poverty<sup>6</sup>, and parents' baseline gender norm attitudes.

### Spillovers

We anticipate that in the case of this specific evaluation we may have “social interactions” spillovers type (as defined by Angelucci and Di Maro 2015) at different spatial levels: classroom, neighborhood and within the household. First, previous evidence from a similar intervention in Brazil highlights that there are very strong spillover effects within classrooms due to social interactions among children and between parents of children in the same class (Bettinger et al., 2020; Lalive and Cattaneo 2009). Since during the school shutdown and the summer period there are no classes, this issue becomes

---

<sup>5</sup> The C4D sample was part of a randomized intervention trial. The treatment had no impacts. As a robustness check, we will also add an indicator for treatment status in the C4D sample.

<sup>6</sup> The exact measurement of this indicator will depend on what is available and comparable in both the C4D and GUP samples. We anticipate it will likely be a household asset index. Alternatively, we consider using household food security as collected in the enrolment call.

less relevant. For the treatment groups for which the nudges continue after classes resume, to the extent that, in our current sample, few parents overlap in terms of having children in the same classrooms, then this issue becomes minor as we randomize at the household level (of course that would be false if there is large overlap). We will track the class and the school the children are attending to have a magnitude of this potential spillover and, if that is large, to estimate the extent to which the treatment has spillovers on peers that were part of the comparison group or that were not part of the experiment. Based on previous evidence on peer effects of interventions that promote educational outcomes (e.g. Lalive and Cattaneo 2009), we expect these spillovers to be positive. Second, there may be neighborhood peer effects on parents' outcomes and children's schooling, such as the ones found in Bobonis and Finan (2009). Specifically, it may be likely that treatment peers (in terms of both parents and children) have influence on their control peers through their behaviors, or may share the information they received in the text. In the case of Bobonis and Finan, effects of those types of social interactions were concentrated among poorest households. Preliminary results from the Cote D'Ivoire trial show that there are positive spillovers for households within a 5-km radius, driven by younger children (early primary grades). We will try to examine spillovers at the neighborhood level as well by directly including in the questionnaire some questions about whether respondents asked for help understanding the SMS, who they asked (teacher, relative, friends, etc.), whether they asked the neighbor, if they discussed the contents of SMS with other parents in the neighborhood etc. We will also ask to the control group if anyone has shared SMS-messages related to children and education with them. Further, we can rely on household GPS coordinates<sup>7</sup> to more precisely estimate spillovers at the neighborhood level. Third, there may be spillovers within the household itself (which is intended as the intervention targets all children in the household). We have included this possibility directly in our research design by sampling two child per household of the target compulsory school age.

To analyze spillover effects empirically at the village level, we will use GPS data and examine the share of treated households within (i) 5 km, (ii) 5-10 km, (iii) 10+ km from each household as a predictor of child schooling and learning outcomes. This allows us to examine the share of households in proximity to each household as a predictor of outcomes.

---

<sup>7</sup> Available for GUP sample, in the case of C4D this has to be confirmed.

## REFERENCES

Angrist, N., Bergman, P., Evans, D. K., Hares, S., Jukes, M. C., & Letsomo, T. (2020). Principles for Phone-Based Assessments of Learning. Working Paper.

Attanasio, Orazio and Kaufmann, Katja, (2014), Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender, *Journal of Development Economics*, **109**, issue C, p. 203-216.

Aurino, E., Gelli, A., Adamba, C., Osei-Akoto, I., & Alderman, H. (2018). *Food for thought? Experimental evidence on the learning impacts of a large-scale school feeding program in Ghana* (Vol. 1782). Intl Food Policy Res Inst.

Bandura A., Barbaranelli C., & Caprara, G.V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, 72, 187-206.

Banerji, R., Berry, J., and Shotland, M. (2017). The Impact of Maternal Literacy and Participation Programs: Evidence from a Randomized Evaluation in India. *American Economic Journal: Applied Economics*.

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

Bettinger, E., Cunha, N., Lichand, G., & Madeira, R. (2020). "Are Effects of Informational Interventions Driven by Salience?". Working Paper.

Doss, C., Fahle, E. M., Loeb, S., & York, B. N. (2017). Supporting Parenting through Differentiated and Personalized Text-Messaging: Testing Effects on Learning during Kindergarten. CEPA Working Paper No. 16-18. *Stanford Center for Education Policy Analysis*.

Favara, M. (2017). Do Dreams Come True? Aspirations and Educational Attainments of Ethiopian Boys and Girls. *Journal of African Economies*.

Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry*. 1997 Jul;38(5):581-6. doi: 10.1111/j.1469-7610.1997.tb01545.x. PMID: 9255702.

Kessler, R.C., Andrews, G., Colpe, .et al (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, 32, 959-956.



Lichand, G., & Wolf, S. (2020). “*Are Parenting Interventions Transferable Across Settings? Evaluating Key Constraints in Sub-Saharan Africa*”. Working Paper.

UNICEF. (2014). The formative years: UNICEF’s work on measuring early childhood development.

Retrieved from :

[http://data.unicef.org/corecode/uploads/document6/uploaded\\_pdfs/corecode/Measuring-ECD-Brochure-HR-10\\_8\\_116.pdf](http://data.unicef.org/corecode/uploads/document6/uploaded_pdfs/corecode/Measuring-ECD-Brochure-HR-10_8_116.pdf)

U.S. Department of Education. (2015b). Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011): User’s manual for the ECLS-K:2011 kindergarten–first grade data file and electronic codebook, public version (NCES 2015-078). Washington, DC: National Center for Education Statistics.

Fertility Behavior and Psychological Stress: The Mediating Influence of Gender Norm Beliefs among Egyptian Women. *Psychology of Women Quarterly*, 25 (2001)

Waszak, C., Severy, L., J., Kafafi, L., Badawi, I. (2001). Fertility Behavior and Psychological Stress: The Mediating Influence of Gender Norm Beliefs among Egyptian Women. *Psychology of Women Quarterly*, 25 (2001)

Wolf, S., Aber, J. L., Behrman, J. R., & Tsinigo, E. (2019). Experimental impacts of the “Quality Preschool for Ghana” interventions on teacher professional well-being, classroom quality, and children’s school readiness. *Journal of Research on Educational Effectiveness*, 12(1), 10-37.