Policy Research Working Paper 9492

Texting Parents about Early Child Development

Behavioral Changes and Unintended Social Effects

Oscar Barrera
Karen Macours
Patrick Premand
Renos Vakis

WORLD BANK GROUP
Poverty and Equity Global Practice
&
Development Impact Evaluation Group
December 2020

Policy Research Working Paper 9492

Abstract

Parenting interventions have the potential to improve early childhood development. Text messages are considered a promising channel to deliver parenting information at large scale. This paper tests whether sending text messages about parenting practices impacts early childhood development. Households in rural Nicaragua were randomly assigned to receive messages about nutrition, health, stimulation, or the home environment. The intervention led to significant changes in self-reported parenting practices. However, it did not translate into improvements in children's cognitive development. When local opinion leaders were randomly exposed to the same text message intervention, parental investments declined and children's outcomes deteriorated. Since interactions between parents and leaders about child development also decreased, the negative effects may have resulted from a crowding-out of some local leaders.

This paper is a product of the Poverty and Equity Global Practice and the Development Impact Evaluation Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at odbarrerar@psemail.eu; karen.macours@psemail.eu; ppremand@worldbank.org; and rvakis@worldbank.org.

Texting Parents about Early Child Development: Behavioral Changes and Unintended Social Effects*

Oscar Barrera (Paris School of Economics)
Karen Macours (Paris School of Economics, INRAE)
Patrick Premand (World Bank)
Renos Vakis (World Bank)

JEL classification: H23, I15, J13, O15.

Keywords: Early Childhood Development, Parenting, Text Messages, Behavioral Change, Field Experiment.

*This study is part of technical and financial support to the Ministerio de la Familia in Nicaragua (Mifan) to design, implement and evaluate interventions to promote parental investments for early childhood development via mobile phones. We are grateful to the project team at Mifan for their guidance and collaboration during the design of this work, as well as the Centro de Investigación de Estudios Rurales y Urbanos de Nicaragua (in particular Veronica Aguilera and Enoe Moncada) for excellent data collection. We are also grateful to Cronos Ltd. for their support with the text messaging platform as well as Telefonica and Movistar for providing low cost mobile phones. Financial support for this research was received from the Government of the Netherlands through the BNPP program and EUR grant ANR-17-EURE-0001. We are grateful for comments received at presentations at the British Academy and PSE, and from Julia Cage, Vera Chiodi, Flore Gubert, and an anonymous DIME reviewer. The computational reproducibility of the results has been verified by DIME analytics. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. All errors and omissions are the authors'. Contacts: odbarrerar@psemail.eu; karen.macours@psemail.eu; ppremand@worldbank.org; rvakis@worldbank.org.

1 Introduction

An estimated 250 million children under 5 years old are at risk of not reaching their developmental potential in the developing world (Black et al., 2017). Early childhood development (ECD) is an important predictor of success later in life (Almond et al., 2018). Early childhood interventions can have high returns given the malleability of the brain (Knudsen et al., 2006) and the hypothesized dynamic complementarities in human capital investments (Heckman, 2006). A growing literature provides experimental evidence on the effects of ECD interventions in the short term (on cognitive or socio-emotional outcomes) or the long term (on school performance, wages, or criminality). The issue is high on the policy agenda in developed (Currie, 2001; Schweinhart, 2005; Nores et al., 2005) and developing countries (Berlinski and Schady, 2015).

Interventions that aim to improve parenting knowledge or skills have been effective at changing parenting practices, for instance through home visiting programs (Jeong et al., 2018). Whether such interventions can be scaled up to improve child development in cost-effective ways remains an open question. Technology-assisted interventions are often considered as a promising avenue for scale-up (Hall and Bierman, 2015). Text-messaging interventions could be an appealing channel to deliver information on ECD due to their low cost, the increasingly widespread use of mobile phones, and their potential to circumvent quality issues with decentralized service delivery (Carta et al., 2013; Ajzenman and López Bóo, 2019). Text messaging interventions are particularly relevant in contexts where personal interactions are not possible, either because of social distancing requirements or more general constraints to accessibility.

A growing literature considers the effectiveness of scaling up early childhood interventions by relying on local service providers, such as community health workers (Hamadani et al., 2019) or social protection agents (Attanasio et al., 2020; Premand and Barry, 2020). The broader economics literature provides evidence of the potential role of local influencers (Banerjee et al., 2019; Alatas et al., 2019) and the existence of social multiplier effects in health or education interventions (Oster and Thornton, 2012; Bobonis and Finan, 2009). In a rural area of Nicaragua close to the setting of this paper, social interac-

tions with local leaders were found to increase the impacts of a conditional cash transfer program on households' investments in education and nutrition, both in the short and medium term (Macours and Vakis, 2014, 2017). Yet a recent literature also highlights negative effects of outside interventions that crowd-out local service providers (Deserranno et al., 2020) or interfere with local institutions (Baldwin et al., 2020).

This paper provides novel experimental evidence on a text message intervention aiming to improve parenting practices in rural Nicaragua. To the best of our knowledge, this is the first paper analyzing the causal effect of providing parenting information by text message on early childhood development in a developing country context. Our first contribution is to test the direct effect of sending text messages to parents on parenting practices and children's cognitive development. Our second contribution is to test whether impacts vary depending on local opinion leaders' exposure to the text messages. We measure impacts on young children's cognitive and socio-emotional development using a series of age-appropriate tests. We also measure impacts on the parental practices, beliefs and investments that were promoted by the intervention. This captures intermediary mechanisms through which the text messages intended to affect ECD outcomes.

The intervention sent daily text messages with advice about parenting practices conducive to early childhood development. The messages covered key risk factors for ECD (Walker et al., 2007b) related to nutrition, health, stimulation and the home environment. They were delivered to randomly selected parents of children between 0 and 6 years of age in 97 villages over a period of approximately 10 months. The text messages were personalized and gave age-appropriate advice. In a random subset of villages, local opinion leaders (such as community health volunteers, preschool teachers, etc.) who also had children received similar text messages. Additional experimental variations were introduced to analyze to what extent impacts differ depending on the individual recipient within the household (mother, father, or both) or the specific risk factors on which the messages focused (nutrition and health, stimulation and the home environment, or both).

Results show that exposure to text messages led to an improvement in reported parenting practices. Parents receiving messages about nutrition and health report better nutritional and hygiene practices. Likewise, parents receiving messages about stimulation

report higher stimulation and progressive beliefs about ECD consistent with those messages. However, no effects on children's cognitive or socio-emotional outcomes are found. Reported changes in parenting practices hence were not sufficient to induce gains in cognitive or socio-emotional development, possibly because they were too small or did not reflect actual shifts in parental investment, or because they were offset by other reactions to the intervention.

To help understand this limited impact of the text message intervention on final outcomes, the analysis considers how impacts vary depending on local opinion leaders' exposure to the text messages. Results show a significant negative impact of leaders' exposure to the intervention on cognitive development of children in their village, a finding that goes against our prior when designing the experiment and earlier results in the literature. Leaders' exposure to text messages also reduces a number of intermediary outcomes, including stimulation and nutritional practices. While the experimental design does not allow us to isolate the exact reasons for these negative effects, we discuss three main potential mechanisms. Results show that the negative effects are unlikely to be driven by confusion about the messages. The evidence points instead to a crowding-out of local opinion leaders and possibly even a boycott. In particular, interactions about ECD practices between local leaders and parents decrease in the randomly-selected villages where leaders received the text messages. We speculate that the intervention may have been perceived by local opinion leaders as interfering with their traditional domain of influence, either demotivating them or possibly leading them to try to offset the messages. While we cannot fully disentangle these potential pathways, the negative effects from leaders' exposure to the text messages and the overall lack of impacts on children's cognitive outcomes are robust. This suggests that the roll-out of technology-based solutions requires careful attention to the social spill-overs and response by local leaders they may generate.

The paper contributes to several strands of the literature. Few studies have assessed the effect of providing parenting advice via text messages on children's cognitive outcomes in early childhood. A set of recent studies on (pre-)kindergarten children in the United States shows that sending text messages to parents can help increase early read-

ing skills (York et al., 2018). Effects vary based on the content, personalization and frequency of the text messages (Cortes et al., 2018; Doss et al., 2019). In an earlier study, Hurwitz et al. (2015) find a positive effect of delivering parental advice via text message on parental engagement in learning activities in early childhood. For older school-age children, Bettinger et al. (2020) find a relatively large impact of text nudges on learning in Brazil.

This paper complements a broader literature showing the potential of text messages as nudges for health-related behaviors such as prevention and disease management (Cole-Lewis and Kershaw, 2010), treatment adherence (Pop-Eleches et al., 2011), child eating behavior (Chai et al., 2019), or vaccination (Kazi et al., 2019). It links to evidence of text messages influencing pro-social behaviors (Konrath et al., 2015). The analysis of the effects of sending text messages to local opinion leaders may also be relevant for the literature on the use of technology to support teachers in Africa and Asia (Gaskell and Mills, 2009; Valk et al., 2010; Walsh et al., 2013), with positive effects found on children's literacy and classroom practices in Kenya (Jukes et al., 2017).

The paper further adds to the literature analyzing impacts of ECD interventions on children's outcomes in developing countries. While no study currently analyzes the impact of parenting advice sent by text message on children's cognitive development, many interventions expose parents to parenting advice through other - often more intensive and direct - means. A seminal study in Jamaica shows that children randomly assigned to receive early childhood stimulation through home visits exhibit large improvements in cognitive outcomes in the short term, and some effects remain twenty years after the intervention (Grantham-McGregor et al., 1991; Walker et al., 2007a, 2000; Powell et al., 2004). Experimental evidence from similar interventions, now available for at least 11 low and middle income countries, generally confirms positive impacts on early childhood development, though effect sizes vary substantially, and delivering effective interventions at scale has proved challenging (Jeong et al., 2018; Attanasio et al., 2020; J-PAL, 2020; Arriagada et al., 2018). In a similar context as the current study, Macours et al. (2015) find positive but small effects of a home visit parenting program.

A range of other interventions have been considered to improve ECD. Cash transfer

programs aim to promote human capital investments in poor households. They have led to shifts in parental investment and gains in cognitive development in Nicaragua (Macours et al., 2012), as well as other developing countries (Paxson and Schady, 2010; Levere et al., 2016). Evidence on long run impacts of exposure in early childhood is also emerging (see Millán et al. (2019) for a review). Experimental studies from interventions primarily focused on behavioral change in nutrition or preventive health practices also show evidence of positive changes in parental investment and nutrition behaviors,[1] but rarely measure impacts on cognition (Fitzsimons et al., 2016; Ahmed et al., 2019; Field and Maffioli, 2020). Recent exceptions include Levere et al. (2016) and Premand and Barry (2020), who find positive effects of behavioral change promotion on parenting practices, but not on cognition.[2]

The rest of the paper proceeds as follows. In the next section, we describe the text message intervention and the data, including ECD measurement. Section 3 presents the empirical strategy. Section 4 presents the main experimental results. Section 5 analyzes potential mechanisms behind the negative impacts of leaders' exposure to text messages. The last section concludes.

2 Intervention, Study Design and Data

2.1 Experimental Design

The text message intervention was implemented from late 2014 to late 2015 in four municipalities in rural Nicaragua.[3] Participants received a daily text message with a short practical recommendation about parenting practices. The messages covered early stimulation, the home environment, nutrition, and preventive health care. The messages were

[1]There is a related literature on direct nutrition interventions, for instance providing supplements. A well-known study in Guatemala shows positive impacts of a nutrition intervention on reading comprehension and test scores years later, and ultimately also higher wages (Maluccio et al., 2009; Hoddinott et al., 2008).

[2]Other ECD interventions do not directly target parental investments. For instance, there is growing but mixed evidence from experimental studies analyzing the impact of preschools in developing countries (Martinez et al., 2017; Bouguen et al., 2018; Dean and Jayachandran, 2019; Blimpo et al., 2019). Ozler et al. (2018) show positive effects of combining preschool teacher training with parenting support.

[3]Totogalpa, Telpaneca, Yalaguina and Palacaguina in the North West of Nicaragua.

sent to parents and focused on the youngest child below 6 years old in the household. The text messages were personalized (making a reference to the name of the child) and age appropriate (with different messages sent to parents of children 0-6 months old, 7-12 months old, 13-24 months old, and older than two years old). Messages were designed to provide actionable advice based on a government early childhood development curriculum, and built on prior experiences with early childhood development programs in the country. The text message intervention was designed by a Nicaraguan early childhood specialist, in consultation with the ministry of the family and with support from the World Bank.

Prior to program implementation, a baseline survey was conducted in 97 rural villages in the four municipalities, including all households with children 0 to 6 years of age or pregnant women. Earlier work in nearby municipalities (Macours and Vakis, 2014) had demonstrated the role of local leaders in influencing behavioral changes for investment in nutrition and health of young children. The baseline survey therefore also identified, among households with children 0-6 (or pregnant women), all households with local opinion leaders regarding early childhood practices. This included community health workers, preschool teachers, primary school teachers, and a village leader.[4] In the 97 villages, there were a total 2,990 households with at least one child younger than 6 years old, including 406 opinion leaders. Five villages had no opinion leaders among households with young children.

For households without leaders, a household-level randomization determined which households would receive text messages: 75% were randomly selected to receive text messages, with the remaining 25% serving as control.

In addition to treatment assignment, orthogonal treatment variations were introduced among households selected to participate in the program. First, we randomly selected whether text messages were sent to (i) the father figure (25 percent of the sample), (ii) mother figure (25 percent of the sample) or (iii) both (25 percent of the sample). Second, the thematic content of the text messages varied. Treated households were randomly

[4]Not included were religious leaders, and those responsible for programs unrelated to health, education or ECD.

assigned to either receive text messages focused on (i) early childhood stimulation and the home environment; (ii) nutrition and preventive health care; or (iii) a combination of topics.[5] The household level randomization was stratified on the level of education of the main caregiver (in most cases the mother of the child) distinguishing those with less than 4 grades, 4-6 grades and more than primary; and on whether there is a male caregiver (typically the father) in the household.[6]

Among the 92 villages with at least one opinion leader, a village-level randomization determined whether household with opinion leaders received no message (27 villages), whether both the men and women in the leader households received messages (27 villages), or whether only the women (27 villages) or only the men (11 villages) in the leader household received text messages.[7] The village-level randomization of the "leader" treatment was stratified on the average level of education of the main caregiver in households with leaders (specifically a variable capturing the tercile in the education distribution).[8]

A registration assembly was organized in each village in December 2014. All households from the treatment and control groups were invited. They all received a basic cell phone labeled as being assigned to the mother or the father of the target child (as per the randomization). This was done because a relatively large share of parents did not have cell phones. Households randomly assigned to have both the mother and father figure receive text messages were given two cell phones. Cell phones were also distributed to control households in order to rule out any effect resulting from an increase in cell phone access. Households assigned to treatment were then assisted to activate the text message intervention. This was done by sending a text message to a central server, with only treated cell phones able to register. Once activated, cell phones started receiving daily messages. Treatment households that did not participate in the assemblies received the

[5]Table A1 provides examples of text messages sent to parents of children 7-12 months old. Note that parents who received both stimulation and nutrition messages did not receive more messages. They also received one message daily, but the theme of the messages varied.

[6]Figure A2 illustrates the household-level randomization design. Note that the number of households in each treatment variation varies slightly due to stratification.

[7]Figure A3 illustrates the village-level randomization design.

[8]Randomization of the content of the messages among leader households was also stratified on access to electricity at home. This is only done for leaders because access to electricity is highly correlated with education among non-leaders.

cell phones, information about the program and help with activation at their home after the assembly. The registration assemblies took a month to complete for all villages, resulting in a take-up (activation) rate of 90%. The intervention cost approximately 50 US$ per household.[9]

2.2 Incentives and Compliance

In addition to receiving daily text messages, registered households had the opportunity to participate in a weekly quiz. They were sent a multiple choice question about one of the parental practices highlighted in earlier messages, and given 24 hours to respond. All individuals who responded correctly entered a lottery to earn free airtime. Every week about 20% of correct answers received 1 US$ airtime, and once a month, winners received 5 US$ airtime (equivalent to 1-2 daily rural wages). The quizzes provided a small incentive for participants to continue reading the text messages. This was considered particularly important as cell phone coverage and electricity access are not universal in the study villages, and receiving text messages at times required walking to a location with signal or charging phones outside of the home. We use the weekly information about households' response to the quiz as a proxy for active participation in the text message intervention.

Figure 1 shows the trivia response rate over time. The response rate peaked at 70 percent about two months after the start of the intervention (i.e. when all registrations were finalized), and then slowly decreased over the year to reach about 25 percent after 10 months.[10] Qualitative field work indicated that this decline reflected three factors. First, some households reported having lost or broken the cellphone or its charger.[11] Second, some households voluntary stopped using the mobile phone all together (some of the phones were seen being used as toys instead). Third, others kept using the mobile phone

[9]Total costs included cell phones (50,000 US$), text messaging distribution platform and incentives (65,000 US$) and field implementation (including registration assemblies) and monitoring (USD 35,000 US$). Given the relatively large fixed costs, costs would be substantially lower at scale.

[10]This percentage is computed over all the cell phone numbers distributed to treatment households, including leaders and non-leaders. Cell phone numbers distributed to control households did not receive the quizzes. The dip in responses in April likely corresponds to a temporary drop in cell phone coverage, which can depend on weather conditions and the cell phone tower signals.

[11]Cell phones for which issues were reported in the first months of the intervention were replaced.

but simply stopped responding to quizzes out of lack of interest or motivation to search for phone signal. Figure A1 shows that the number of treated individuals participating in weekly quizzes evolved similarly for the three thematic variations.

2.3 Data

Baseline data were collected in the fall of 2014 using a short household survey instrument. It captured information on household composition, socio-economic status, cell phone usage, social interactions and economic activities, as well as child-specific information regarding health practices, preschool participation, nutritional intake, early childhood practices and attitudes. The baseline was used to identify for each child a "mother" and a "father" figure, i.e. the child's biological mother and father, or in their absence, the individuals with similar caregiver roles, as defined by the household.

A follow-up survey was implemented in July and August 2015, approximately nine months after households had started receiving the text messages. The duration of exposure is similar to other early childhood interventions where impacts on ECD outcomes were found in nearby municipalities in Nicaragua (Macours et al., 2012).[12] The follow-up survey targeted children in sample households, as well as the main caregiver. We limit the analysis to children under 7 years of age at follow-up (2,803 children), in order to use the full battery of tests, as detailed below. Questions on early childhood practices and attitudes were asked separately to mother and father figures.

The child-level follow-up instrument includes a number of tests to assess early childhood development. First, the four sub-scales of the Denver Developmental Screening Test (Frankenburg and Dodds (1967)) were used to assess social-personal, language, fine motor, and gross motor skills for all children between 12 and 84 months of age. We use a modified version of the Denver previously used as part of a national early childhood stimulation program in Nicaragua.

Three additional tests were conducted for children between 36 and 84 months old: 1)

[12]The follow-up survey was fielded a bit before the end of the text message intervention so that data collection could take place before the rainy season, when accessibility can become challenging. That said, about 2 percent of observations was collected later, during a short tracking phase in May 2016 when missing households were re-visited.

The TVIP, the Spanish-speaking version of the Peabody Picture Vocabulary Test (PPVT) (Dunn et al., 2006); 2) The digit span, an associative numeric memory test; and 3) A version of the marshmallow test to measure self-control.[13] The test was adapted to the local (humid) survey context by using hard red candy instead of a marshmallow.

The tests have been applied to similar populations in Latin America, including in previous studies in Nicaragua (Macours et al. (2012) and Barham et al. (2013)), Ecuador (Paxson and Schady (2007) and Paxson and Schady (2010)) and Mexico (Fernald et al. (2008)). All tests were extensively piloted in the field and adjusted when necessary. Test administrators were selected based on their background (trained as psychologists, social workers, or similar fields) and for their ability to quickly establish a rapport with young children. They were intensively trained on the standardized application of the tests, as well as on putting children at ease before starting the tests. Tests were administered at home. The privacy of the test-taker and the confidentiality of the results were ensured throughout the process. Test administrators were randomly assigned to households. The quality and standardized application of the tests was closely monitored in the field. A key advantage of the tests is that they provide observed and objective measures of child development, rather than parent-reported measures that may suffer from reporting biases.[14]

The test scores are aggregated using principal component analysis to determine factor weights in the control group. We retain the first principal component as a summary indicator of early childhood cognitive development. The first component accounts for 32% of the overall variation in the 7 tests, and for 35% of the overall variation in the 4 subcomponents of the Denver. Table A2 reports the factor weights, showing that cognitive development tests all have relatively high weights (Denver language, TVIP, memory, fine motor skills).

The aggregate cognitive score constitutes our main ECD outcome. In addition, socioemotional skills are measured using the Strength and Difficulties questionnaire (SDQ) and a behavioral screening test consisting of questions to caregivers about a set of positive and

[13]The marshmallow test is a well-known test of delayed gratification, evaluating preschool-age children's ability to wait before eating a marshmallow in exchange for being rewarded an additional one. Mischel et al. (1989) show that the test predicts outcomes later in life.

[14]Only a few items of the Denver are reported by caregivers.

negative behaviors for children 36 to 83 months old. These scales were complemented with questions on inhibitory control and positive demeanor. Another standardized set of age-appropriate behaviors was asked for children aged 18 to 35 months old. We consider socio-emotional measures as a secondary outcome. We analyze it separately because information self-reported by caregivers is sensitive to potential reporting biases, which may be influenced by the text message intervention. As for the cognitive score, we construct an aggregate index across the various socio-emotional development subscales. We use principal component analysis to determine factor weights in the control group (see Table A3).[15]

The text message intervention aimed to improve early childhood cognitive and socio-emotional development by changing parental investment behavior and attitudes linked to ECD risk factors. To measure intermediate outcomes related to parenting practices, we construct seven indices for parental investments in nutrition, protein intake, micro-nutrient intake, stimulation, health, and indices for caregiver attitudes regarding ECD, and hygiene. Each index is constructed using a set of questions about parental behaviors, ECD risk factors, and attitudes.[16] The nutrition index is based on a set of questions measuring the number of days the child receives nutritious food during the week before the survey. The stimulation index aggregates questions on whether the caregiver gives toys to the child, reads or tells stories to the child, and whether there is pen and paper in the house for the child to draw. The health index is based on twenty questions related to preventive health behaviors, including tooth brushing, hand washing, use of mosquito nets, vaccination, boiling water etc. The micro-nutrient index is based on three variables measuring whether the child has received vitamin A, ferrous sulfate, or de-worming medicine during the last six months. The animal protein index measures consumption of proteins (milk, eggs, cheese, meat) during the last week. The attitudes index is based on variables measuring the caregiver's opinion about ECD and stimulation practices. Finally, the hygiene index captures the condition of the child during the interview. In contrast with

[15]While some of the subscales have low or even negative weights, we restrict the analysis to the first principal component as for the cognitive outcome. For comparability, we also restrict the analysis to the sample of children with cognitive tests.

[16]We provide additional details on the questions used in each index when discussing related results in the next section.

other intermediate outcomes, the hygiene index is directly observed by the enumerator. Each index score is standardized using the mean and standard deviation in the control group.

2.4 Balance and Attrition

The endline sample is balanced on baseline characteristics between treatment and control households (Table 1). It is also balanced between villages with and without opinion leaders assigned to treatment (Table 2). Table 1 provides relevant contextual information about the study setting. Both fathers and mothers have low levels of education, with mothers having completed 3.3 years of education on average, and fathers 3 years. 26 percent of household heads are illiterate. 76% of target children live with their father at baseline, and 96% with their mother. Children on average consume meat and vegetables one to two days per week, and fruit and eggs about 3 days per week. By contrast, coffee consumption is high (four days per week). Only 22% of children had an adult read to them in the month before the baseline survey.

Attrition at follow-up is 13.8 percent for tests administered to the child, 10.4 percent for questions asked to the main caregiver, 12 percent for the mother-specific questions and 21 percent for father-specific questions.[17] Attrition is balanced between treatment and control groups for the household and mother-reported questions, but there is a slight imbalance (significant at the 10%) for child-level outcomes (Table A4). Baseline observables are, however, similar for attrited observations in the treatment and control groups for all outcomes (Table A5). This suggests that the profile of children and parents lost due to attrition are similar in the treatment and control groups. This is confirmed by the balance tests shown in Table 1. Moreover, attrition is not significantly different between households receiving different thematic content (Table A4, panel B), and is also balanced between villages with and without opinion leaders assigned to receive the text messages. Indeed, the coefficients for child and household level attrition in panel C of Table A4 are close to zero, suggesting attrition is not likely to bias the estimated effects of leaders'

[17] We do not use the father-specific questions given the relatively higher attrition among fathers.

assignment to the text message intervention.

3 Empirical Specification

We estimate the following child-level intent-to-treat regression:

$$Y_{iv} = \alpha T_i + \beta X_i + \delta_v + \upsilon_{iv} \tag{1}$$

where Y_{iv} is the outcome for child i in village v; T is an indicator denoting treatment assignment, which takes the value of one for children in households randomly assigned to receive the text message intervention; X is a set of control variables, including the stratification variables[18] as well three-monthly age dummies, the gender of the child, and test administrator fixed effects. We also include a binary variable indicating whether data were collected during the tracking phase in 2016.[19] δ_v is a village fixed effect, which controls for the village-level assignment of the leader treatment.

We also run intent-to-treat regressions to estimate the effects of the randomized treatment variations:

$$Y_{iv} = \alpha_1 T_{1i} + \alpha_2 T_{2i} + \alpha_3 T_{3i} + \beta X_i + \delta_v + \upsilon_{iv} \tag{2}$$

where T_1 , T_2 , and T_3 are indicators denoting children whose parents were randomized to receive text messages focused on nutrition, stimulation, or both nutrition and stimulation. We use a similar specification to estimate treatment effects on children's outcomes depending on whether the mother, the father or both parents were randomly selected to receive the text messages.

The household-level randomization design identifies treatment effects net of spill-overs between households within villages.[20] To measure spill-over effects directly, we

[18]The stratification variables include: the level of education of the main caregiver, whether there is a male caregiver in the household, whether it is a leader household, and whether it is a leader household with access to electricity.

[19]As indicated above, in May 2016 the data collection team tracked 63 households that were not found during the initial visit. The control variable accounts for the fact that the intervention had ended by then, and for the age-sensitivity of the tests.

[20]While positive spill-overs between households because of information sharing could lead to an under-

focus on spill-over effects from opinion leaders' exposure to the text message intervention through the following model:

$$Y_{iv} = \beta_1 L_v + \delta T_i + \beta_2 X_i + \epsilon_{iv} \tag{3}$$

Y_{iv} is the outcome for child i in village v. L_v takes the value of 1 in villages where opinion leaders were randomly assigned to receive text messages. Controls include the stratification variables used for the leader randomization (the level of education of the leader) and other household-level controls as in specification 1. To isolate spillover effects, equation (3) is estimated by excluding the households of the opinion leaders themselves. The coefficient of interest (β_1) captures the spill-over effects of being in a village where leaders receive text messages. Randomization took place at the village level and we cluster standard errors at the village level.

To avoid concerns about multiple hypothesis testing, we focus only on two main outcomes. Cognitive development is the primary outcome, and socio-emotional development the secondary outcome of interest. As discussed in Section 2.3, we create an index for each outcome by aggregating related measures. We also test for impacts on intermediate outcomes capturing parenting practices, which are also aggregated by domain.

4 Results

4.1 Direct Effects of Text Message Intervention

The main results on the effect of the text message intervention on early childhood development are reported in Table 3. Panel A reports ITT estimates for the aggregate cognitive development index for all children age 1-7 years at endline (column 1), as well as separately for children age 1-3 years (column 2), and 3-7 years (column 3). The estimated impact on early childhood development is very small and not significantly different from

estimation of impacts, it likely cannot explain why estimates show impacts on intermediate outcomes but not on final ECD outcomes.

zero.[21] This result holds for the different age groups. Table A6 in the appendix shows that it also holds for the individual tests when they are analyzed separately. Table A7 shows consistent results across the treatment variations focusing on nutrition, stimulation or both, as well as depending on whether the intervention targets the mother, the father, or both. Tables A8 and A9 further show that there is no effect on the overall behavioral (socio-emotional) index (column 1), which holds for most of its individual components as well (in additional columns).

Panel B of Table 3 reports ITT estimates by level of education of the main caregiver. This shows a small negative coefficient of -0.12 standard deviation among children of the least educated caregivers (those with 3 or less years of education), significant at the 5% level (column 1). In contrast, interaction terms are positive for children whose caregiver completed primary school. But this does not result in an overall positive effects, except for the youngest children (significant at the 10% level with an effect size of 0.16 standard deviation). Finding more variation among children aged between 1 and 3 years old is in line with the literature suggesting higher malleability at a young age. It is also consistent with the hypothesis that interventions in early childhood can be particularly important for cognitive development, a key rationale to target this age group. Differences by education levels suggest that caregivers' ability to read and understand the messages could (intuitively) be important for the messages to be effective. There is no such heterogeneity, however, for older children for whom there is little impact overall.

Table 4 shows results for intermediate outcomes for all children between 12 and 83 months old. Positive ITT estimates are found for all intermediate outcomes, with small magnitudes between 0.06 to 0.16 standard deviation. Results for individual questions used to construct the indices are reported in the Appendix Tables A10, A11, and A12. Panel B of Table 4 reports estimated effects on the 7 intermediate outcomes, separately for each (randomized) variation in the content of messages. The impacts on the indices for nutrition, micro-nutrients intake, proteins and hygiene are larger for households assigned to the nutrition and preventive health messages (though not significantly so). Impacts

[21] Given the small and insignificant effect of direct exposure to text messages on cognitive outcomes, using bounds to correct for the small attrition differential mentioned in section 2.4 would not change the findings.

on stimulation and caregivers' attitudes toward ECD are significantly larger for those assigned to the stimulation and home environment messages. Overall these results show that changes in reported parental investments and practices broadly reflect the content of the text messages that caregivers received. That said, none of the thematic variations lead to significant improvements in early childhood development outcomes (as mentioned above in relation to Table A7).

Since most of the intermediate outcomes measure parenting behavior as reported by the child's main caregiver, it is possible that these results capture parents increased knowledge and awareness about these practices, rather than actual shifts in behavior (e.g. if social desirability bias makes parents that had been exposed to text messages feel more compelled to say they implement certain practices even if they do not). Importantly, however, the hygiene index, which is based on direct observations by the enumerators, also shows a significant improvement. Finding positive results on intermediate outcomes but not on final ones could also simply be because the shifts in behavior were too small to translate into cognitive or socio-emotional gains. Or it could be that the potential gains from modest shifts in parental investments as a result of direct exposure to the text messages were offset by other reactions to the intervention, to which we turn next.

4.2 Impact of Opinion Leaders' Exposure to Text Message Intervention

Table 5 reports the estimated β_1 coefficients from equation 3 on the main outcome (as in Table 3). Panel A shows the estimated effect of living in a village where opinion leaders were sent the text messages.[22] It shows that sending messages to opinion leaders has a negative spillover effects on cognitive outcomes of children from other (nonleader) households in the same village. Children living in villages where the leader was assigned to treatment have a score on average 0.11 standard deviation lower than children from households in villages where leaders were not treated (column 1). While the spillover coefficient is very small and insignificant for the younger children, it is negative and highly significant for children between 3 and 7 years old (-0.14 standard deviation). When

[22]Recall that we only consider social interaction effects with opinion leaders who themselves have young children. The intervention delivered text messages that were personalized about recipient's own children.

considering impacts on individual tests, Table A13 shows that the negative results are strongest for the two language scores (Denver language and the receptive vocabulary test), which are often considered the best proxies for cognition. On the other hand, Tables A8 and A9 show no significant impact of the leader treatment on socio-emotional outcomes.

Panel B of Table 5 reports the leader spillover effects on cognitive outcomes distinguishing by caregivers' level of education. Results show a negative spillover effect from leaders' exposure to text messages for all age groups among children whose caregivers have fewer than 4 years of schooling. Children with the least educated parents hence appear the most affected by the negative leader effect. That said, for the older children, and for all children together, the effect is also negative for those whose parents completed primary education (p-value for the joint significance test for the high educated group is 0.05 and 0.03). Similarly when distinguishing by education levels of the leaders, Table A14 shows that the negative effects are found for leaders with the lowest and the highest education levels.

Exposure of opinion leaders to parenting text messages leads to a deterioration of early childhood outcomes among children in their village. As such, the effects of opinion leaders' exposure to text messages go in the opposite direction than anticipated, and suggest a possible negative influence of these opinion leaders. Before delving into possible mechanisms, we first use an alternative specification to test the robustness of this unanticipated effect. We define a variable measuring the number of dwellings between each household and the closest opinion leader. The variable captures physical distance to opinion leaders, which can be used to test whether the negative leader effects are driven by households that are closer to the leaders. Of course, the variable could also capture remoteness more generally (if opinion leaders live in more central locations), or social distance (if, for instance, members of the same extended families live closer to each other), among other factors. Even so, the interaction effect between the distance variable and the leader treatment provides a useful check about the plausibility that the negative experimental leader effects comes from exposure to those opinion leaders. The interaction effect in Table 6 (panel A) shows that the negative leader effect is indeed stronger for households living

close to the leaders, and weakens as distance increases (column 1). This effect is particularly strong for younger children (column 2). For them, the significant interaction term indicates that the negative leader spillover effect disappears for households living 6 or more dwellings away from the leader's house.

Panel B further shows that the negative leader effect is particularly strong for children from households that do not receive text messages themselves (-.25 standard deviation for the older children). Possibly this is because it was not offset by any positive direct treatment effect. For older children, the interaction between the household-level treatment and leader spillover is positive and significant at 10%, suggesting that negative effects of leaders' exposure may be weaker for those receiving text messages. Nevertheless, the joint significance test reported at the bottom of the table shows that the negative leader effect is also significant for households who received the text messages. For younger children, results suggest no significant difference in the leader effects between treatment and control households.

Table 7 (panel A) reports estimated coefficients from equation 3 for the seven intermediate outcome indices. Negative leader effects are found for several of the indices and related ECD risk factors, specifically in the domains of nutrition, micronutrients and proteins (see columns 1, 4, 5). This highlights a likely channel through which leaders' exposure can negatively influence parenting practices and contribute to negative impacts on ECD outcomes. Table 7 (panel B) reports results on the same outcomes by the caregivers' education level but shows no clear pattern of heterogeneity.[23]

[23]Appendix Tables A15 and A16 further test whether the negative leader effects are stronger for households who had more social interactions about parenting practices at baseline. At baseline, households were asked whether they had talked to different types of members in the community in the last 7 days about ECD practices. We measured in particular whether anybody in the household had any interaction about ECD practices with the health promoter, pre-school teacher, primary school teacher, other teachers, family members, neighbors, or elected leaders in the village. About half of the households had talked to at least one other community member about ECD practices. Results in Tables A15 and A16 show that there is no clear heterogeneity by this indicator of baseline social interactions.

5 Mechanisms

The negative effect of opinion leaders' exposure to the text message intervention on young children's development outcomes and (some) parental practices goes against the initial hypothesis motivating this experimental variation. The experimental design does not allow to causally identify the underlying mechanism. Still, given its potential relevance for the design of similar information interventions, we attempt to further unpack this result, acknowledging that the analysis is more speculative in nature.

This section considers a number of possible explanations. First, confusion: because of the design, leaders received a different text message than many of the other parents close to them. Potentially, this could have led to confusion when they exchanged information. Second, boycott: leaders' prior beliefs about optimal parenting practices may not have coincided with the text messages, and leaders receiving the messages may have attempted to offset them. And third, crowding-out: local leaders may have felt de-legitimized or demotivated by the text message intervention, which directly interfered with their sphere of influence. Or parents might have decreased their interactions with leaders now that they obtained information directly through text messages.[24]

5.1 Confusion

Within each village, households received various types of text messages. This is because of the randomized variation in content, but also because different age-appropriate mes-

[24] Another mechanism may be competition, whereby leaders receiving text messages favored advancing their own children's outcomes at the expense of other children. For instance, leaders may have reallocated resources (such as time) to their own children at the cost of their responsibilities in the community. To analyze whether such competitive motives contribute to the observed negative effects, we analyze treatment effects on leaders' children. In absence of the intervention, leaders' children have better ECD outcomes (Table A20). This is in line with the higher education levels and general status of opinion leaders in the communities. However, Table A20 shows that there is no significant treatment effect on leaders' children. Table A21 further shows that changes in intermediate outcomes are not significantly different between leader and non-leader households. (Proteins are an exception, with a negative and significant interaction between the treatment and leader). Overall, changes in behavior appear more limited for leaders, possibly because leaders engage more in some of the practices irrespective of the intervention. While these results are an imperfect test of a competition hypothesis, the lack of impact on leaders' practices and their children's outcomes suggest that leaders exposed to text messages did not increase investments in their own children. This makes it unlikely that the negative effect on non-leaders' children resulted from a deliberate effort by leaders to favor their own children.

sages were sent to households with children in different age groups. Very often, opinion leaders therefore received a different text message than other households living nearby. For instance, on a given day, an opinion leader may have received a text message about nutrition of her 12 month old child, while a household in her proximity received a message on stimulation of her 4 year old child. To the extent that opinion leaders engage with households in their villages based on the messages they receive, this could have caused confusion, possibly especially among the least educated parents.

We investigate this hypothesis by analyzing whether ECD outcomes are better in cases when parents and the closest leader received the same message (compared to those who received a different message). The point estimate of the difference between those two groups is very small, -0.00 (s.e. 0.09), hence providing no evidence in support of the confusion hypothesis. Note, however, that given the large number of potential combinations of text messages between leader and non-leader households, there are relatively few pairs of leader and non-leader households who received the same text message (80 of a total 1,004 pairs), so that the comparison may be underpowered. Even so, the precisely estimated zero suggests that confusion is unlikely to be a major driver of the negative leader results.[25]

5.2 Boycott

Another potential explanation for the negative leader effect could be that opinion leaders have strong prior beliefs about parenting practices, and may tell people not to believe messages if they go against their own beliefs. Qualitative evidence does suggest that some text messages promoted practices diverging from traditional beliefs.[26] The baseline survey includes a set of five questions about caregivers' beliefs regarding early childhood development, in particular on stimulation and the home environment. The baseline questions about beliefs are similar to those used at follow-up to capture attitudes toward early

[25] Also recall that the negative leader effect is observed for children from households that do not receive text messages, as mentioned above.

[26] For example, several households said they had learned they could give eggs to children, while before they thought this would hurt them. Other examples include parents stopping to give coffee to kids, abandoning harsh disciplining practices, or letting girls plays with toys traditionally considered to be for boys.

childhood development. The baseline answers to those questions show that half the leaders disagree with 2 or more of the 5 messages about attitudes toward early childhood development. Table 8 and 9 test for heterogeneity in the leader effect based on the baseline attitudes toward ECD of the closest leader. The interaction term indicates whether the leader effect is different for households living close to leaders whose prior opinions were largely in line with the stimulation and home environment text messages (i.e. they agreed with 4 out of 5 promoted messages, and we hence label them as being "agreeing" leaders). With the exception of the intermediary outcome on micro-nutrients, interaction effects are not significant. For the early childhood development outcomes, they go in the direction of the boycott hypothesis for the younger children, but not for the older ones. Unfortunately, there are no baseline measures of leaders' beliefs about nutrition and health practices, which are the risk factors where the negative leader effects appear to be the strongest (in Table 7).

To complement this analysis, we also explore whether the negative opinion leader effect varies by leaders' predicted commitment or engagement with respect to the text message intervention. We use the frequency of leaders' participation in the quizzes as a proxy for leaders' interest in the text message intervention.[27] Of course, we do not have such a proxy for leaders in the control group. We therefore predict the frequency of quiz participation for leaders randomly assigned to treatment, using a Lasso estimator to select baseline observables.[28] We obtain a relatively precise prediction (with a R^2 of 0.65) and define "committed" leaders as those who are predicted to be in the top 20% of quiz participation. This threshold corresponds to leaders participating approximately half of the time. We can then analyze whether the leader effect differs depending on the predicted commitment of the closest leader to each non-leader household.[29]

Table 10 shows that negative leader effects in column 1 and 3 are attenuated for children from households in proximity of committed leaders. Although the interaction term

[27] By this measure, opinion leaders do not appear to be more committed to the intervention than other households: there is no significant difference in the number of quizzes that leader households responded to (22 quizzes) compared to households without leaders (21 quizzes).

[28] See appendix for details, including a list of covariates used in the prediction model.

[29] We consider heterogeneity based on the closest leader as the prediction is done for each leader separately (and hence varies within a community). The closest leader is defined as in section 4.2.

is not significant, the overall leader effect on ECD outcomes for those children is not significantly different from zero (bottom row of the table). This pattern is confirmed in Table 11, which shows the same specification for the intermediate outcomes. There are no significant leader effects for households in proximity of committed leaders for 6 of the 7 outcomes (the micro-nutrient index is an exception). Moreover the point estimates indicate that the leader effects for households whose closest leader is predicted to be committed to the program are in fact positive for stimulation, health, and attitudes, though none of these effects are significant.[30] While a high predicted program participation may be capturing many different characteristics of the leaders (and those living in their proximity), this heterogeneity analysis provides suggestive evidence that the negative effects of opinion leaders are driven by leaders that were less likely to actively engage with the program. This result could be consistent with the boycott hypothesis. It is also consistent with a de-legitimization or demotivation of leaders inducing a crowding-out of interaction with leaders, a mechanism we now consider in more detail.

5.3 Crowding Out

Another potential mechanism is that the text message intervention crowded out local leaders. On the one hand, the text message intervention may have been perceived by local opinion leaders as demotivating or de-legitimizing them because it directly interfered with their sphere of influence. If this led to less interactions between parents and local opinion leaders, parents may have missed out on personal advice tailored to the specific needs of their children. This may have contributed to the negative effects observed in Table 5. By demotivating the service providers, it may also have lowered the quality of education or basic health services in the community. On the other hand, the direct provision of information to parents may have decreased the need for parents to seek guidance from local leaders through direct interactions with them.

Table 12 shows that leaders' randomized exposure to text messages led to significantly lower interactions about ECD practices between opinions leaders and parents. Impor-

[30] Appendix Table A22 and A23 show results for each quintile of predicted participation.

tantly, there is no effect of direct exposure to the text messages on interactions between caregivers and leaders. This suggests that caregivers themselves did not decrease their consultations with local leaders after receiving parenting information through text message. Instead, the effects of leaders' exposure to the text messages on interactions about ECD interactions is negative for all types of leaders. It is significant for primary school teachers, other educators and community leaders. This is broadly consistent with the results indicating that the negative leader effects were also observed for the most educated leaders. Moreover, a comparison with Table A24 shows that these negative effects on social interactions are limited to opinion leaders that were directly targeted with the text messages. There are no significant reductions in social interactions about ECD practices with family or neighbors, and possibly even a slight increase in interactions with religious leaders. While the data do not allow us to analyze local service delivery or the nature of interactions between leaders and parents in more detail, these results suggest that opinion leaders' exposure to the text messages led to a crowding out of their direct interactions with parents of young children in the community. The evidence is also consistent with the decrease of interaction originating from leaders rather than caregivers.

6 Discussion and Conclusion

A large number of children in developing countries suffer from signs of cognitive delays, which start at a very young age and affect their lifelong prospects. Experimental evidence has shown that interventions aiming to improve parental investments and practices can positively impact cognitive and socio-emotional development in early childhood. Growing evidence has also demonstrated that early investments can lead to longer-term gains in terms of education, earnings and social outcomes. In light of this evidence, there is growing interest in interventions that seek to improve parenting practices, with major questions on whether they can be delivered cost-effectively at scale. Text message interventions offer an appealing alternative to more intensive approaches such as home visiting programs, especially at times when personal interactions are not possible due to limited accessibility or social distancing requirements.

This paper presents results from an RCT of an intervention sending daily text messages on parenting practices to caregivers in poor households in rural Nicaragua. The program enhanced knowledge and improved reported parental practices associated to nutrition and stimulation. However, no improvements were found on children's cognitive or socio-emotional outcomes. These results contrast with those from more intensive interventions improving parental investment and early child development in Nicaragua and elsewhere in the developing world. Importantly, the results are obtained in a context where cell phone coverage and access to electricity is far from universal, where low levels of education may have limited parents' ability to fully internalize the text messages and where few households were covered by a government-led ECD program. These contextual factors can help explain the lack of more positive results, and are worth considering when contemplating the external validity of our findings. At the same time, lack of electricity, limited cell phone ownership, low coverage of ECD programs, and low literacy are common in many parts of the developing world.

The second main result of this paper is that opinion leaders' randomized exposure to the text message intervention led to a significant decline in cognitive outcomes among children from non-leader households. These spill-over effects are stronger for children of the least educated parents, and for those living closest to the local opinion leaders. We explore potential mechanisms that may explain this negative effect of opinion leaders' exposure to text messages. Leaders' exposure to text messages led to a decrease in interactions about ECD practices between leaders and parents, suggesting a possible crowding out effect. Possibly the text messages led to demotivation or de-legitimization of local opinion leaders, and a reduction of their efforts in the provision of ECD information or services.

The results in this paper contrast with those of a CCT program in a nearby region of Nicaragua in two dimensions. First, while both interventions led to significant changes in reported parenting practices, the text messages did not result in improvements in ECD outcomes, while the CCT program did (Macours et al., 2012). Possibly the changes in knowledge and practices resulting from the text messages (3-4 times smaller than those obtained with the CCT) were simply too small to translate into changes in final ECD

outcomes. Second, Macours and Vakis (2014, 2017) show that the CCT impacts were in part induced by an increase in social interactions with local leaders, while this paper shows negative spill-overs and a potential crowding-out of local leaders. While the CCT program gave local leaders an explicit recognized role in the information dissemination and mobilization around the intervention, there was no active role for local leaders in the text message intervention, which can possibly help explain these contrasting findings. The contrasting results also more generally point to a need to better understand how to leverage social interactions to enhance positive behavioral change.

Overall, the lack of impacts of text messages to parents on ECD outcomes and the negative effects of leaders' exposure to the text messages call for caution before advocating for the large-scale roll-out of text message parenting interventions in high-poverty settings. Such interventions may need to be complemented with other types of ECD programs. And they may need to explicitly incorporate strategies to crowd-in local opinion leaders, for instance by empowering them with a specific role during implementation. These hypotheses would deserve to be tested in future research.

References

Ahmed, A., Hoddinott, J., and Roy, S. (2019). Food transfers, cash transfers, behavior change communication and child nutrition: Evidence from Bangladesh. Working Paper No. 1868, Washington, D.C.: International Food Policy Research Institute (IFPRI).

Ajzenman, N. and López Bóo, F. (2019). Lessons from behavioral economics to improve treatment adherence in parenting programs: An application to SMS. IZA Discussion Paper No. 12808.

Alatas, V., Chandrasekhar, A. G., Mobius, M., Olken, B. A., and Paladines, C. (2019). When celebrities speak: A nationwide twitter experiment promoting vaccination in Indonesia. Working Paper 25589, National Bureau of Economic Research.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, 56(4):1360–1446.

Arriagada, A.-M., Perry, J., Rawlings, L., Trias, J., and Zumaeta Aurazo, M. (2018). Promoting early childhood development through combining cash transfers and parenting programs. Policy Research Working Paper No. 9368, Washington, D.C.: World Bank Group.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2020). Estimating the production function for human capital: Results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1):48–85.

Baldwin, K., Karlan, D., Udry, C. R., and Appiah, E. (2020). How political insiders lose out when international aid underperforms: Evidence from a participatory development experiment in Ghana. Working Paper 26930, National Bureau of Economic Research.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*, 86(6):2453–2490.

Barham, T., Macours, K., and Maluccio, J. A. (2013). Boys' cognitive skill formation and physical growth: Long-term experimental evidence on critical ages for early childhood interventions. *American Economic Review*, 103(3):467–71.

Berlinski, S. and Schady, N. (2015). More bang for the buck: investing in early childhood development. In *The Early Years*, pages 149–178. Springer.

Bettinger, E., Cunha, N., Lichand, G., and Madeira, R. (2020). Are the effects of informational interventions driven by salience? Working Paper No. 350, Department of Economics - University of Zurich.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J., et al. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, 389(10064):77–90.

Blimpo, M. P., Amaro Da Costa Luz Carneiro, P. M., Jervis, P., and Pugatch, T. (2019). Improving access and quality in early childhood development programs : Experimental evidence from the Gambia. Policy Research Working Paper No. 8737, The World Bank.

Bobonis, G. J. and Finan, F. (2009). Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics*, 91(4):695–716.

Bouguen, A., Filmer, D., Macours, K., and Naudeau, S. (2018). Preschool and parental response in a second best world evidence from a school construction experiment. *Journal of Human Resources*, 53(2):474–512.

Carta, J. J., Lefever, J. B., Bigelow, K., Borkowski, J., and Warren, S. F. (2013). Randomized trial of a cellular phone-enhanced home visitation parenting intervention. *Pediatrics*, 132(Supplement 2):S167–S173.

Chai, L. K., May, C., Collins, C. E., and Burrows, T. L. (2019). Development of text messages targeting healthy eating for children in the context of parenting partnerships. *Nutrition & Dietetics*, 76(5):515–520.

Cole-Lewis, H. and Kershaw, T. (2010). Text messaging as a tool for behavior change in disease prevention and management. *Epidemiologic Reviews*, 32(1):56–69.

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., and York, B. N. (2018). Too little or too much? Actionable advice in an early-childhood text messaging experiment. *Education Finance and Policy*, pages 1–44.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2):213–238.

Dean, J. T. and Jayachandran, S. (2019). Attending kindergarten improves cognitive but not socioemotional development in India.

Deserranno, E., Nansamba, A., and Qian, N. (2020). Aid crowd-out: The effect of NGOs on government-provided public services. Working Paper No. 26928, National Bureau of Economic Research.

Doss, C., Fahle, E. M., Loeb, S., and York, B. N. (2019). More than just a nudge supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, 54(3):567–603.

Dunn, L. M., Dunn, L., and Arribas, D. (2006). Peabody, test de vocabulario en imágenes. *Madrid: TEA ediciones*.

Fernald, L. C., Gertler, P. J., and Neufeld, L. M. (2008). Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. *The Lancet*, 371(9615):828–837.

Field, E. and Maffioli, E. M. (2020). Are behavioral change interventions needed to make cash transfer programs work for children? Experimental evidence from Myanmar. Mimeo.

Fitzsimons, E., Malde, B., Mesnard, A., and Vera-Hernandez, M. (2016). Nutrition, information and household behavior: Experimental evidence from Malawi. *Journal of Development Economics*, 122:113–126.

Frankenburg, W. K. and Dodds, J. B. (1967). The Denver developmental screening test. *The Journal of Pediatrics*, 71(2):181–191.

Gaskell, A. and Mills, R. (2009). Using mobile technology for learner support in open schooling: A report to the commonwealth of learning.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., and Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study. *The Lancet*, 338(8758):1–5.

Hall, C. M. and Bierman, K. L. (2015). Technology-assisted interventions for parents of young children: Emerging practices, current research, and future directions. *Early Childhood Research Quarterly*, 33:21–32.

Hamadani, J. D., Mehrin, S. F., Tofail, F., Hasan, M. I., Huda, S. N., Baker-Henningham, H., Ridout, D., and Grantham-McGregor, S. (2019). Integrating an early childhood development programme into Bangladeshi primary health-care services: an open-label, cluster-randomised controlled trial. *The Lancet Global Health*, 7(3):e366–e375.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902.

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., and Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The Lancet*, 371(9610):411–416.

Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., and Wartella, E. (2015). Supporting head start parents: impact of a text message intervention on parent-child activity engagement. *Early Child Development and Care*, 185(9):1373–1389.

J-PAL (2020). Encouraging early childhood stimulation from parents and caregivers to improve child development. Technical report.

Jeong, J., Pitchik, H. O., and Yousafzai, A. K. (2018). Stimulation interventions and parenting in low-and middle-income countries: A meta-analysis. *Pediatrics*, 141(4):e20173510.

Jukes, M. C., Turner, E. L., Dubeck, M. M., Halliday, K. E., Inyega, H. N., Wolf, S., Zuilkowski, S. S., and Brooker, S. J. (2017). Improving literacy instruction in Kenya through teacher professional development and text messages support: A cluster randomized trial. *Journal of Research on Educational Effectiveness*, 10(3):449–481.

Kazi, A. M., Ahsan, N., Khan, A., Jamal, S., Kalimuddin, H., Ghulamhussain, N., Wajidali, Z., Muqeet, A., Zaidi, F., Subzlani, M., et al. (2019). Personalized text messages and automated calls for improving vaccine coverage among children in Pakistan: Protocol for a community-based cluster randomized clinical trial. *JMIR Research Protocols*, 8(5):e12851.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., and Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155–10162.

Konrath, S., Falk, E., Fuhrel-Forbis, A., Liu, M., Swain, J., Tolman, R., Cunningham, R., and Walton, M. (2015). Can text messages increase empathy and prosocial behavior? the development and initial validation of text to connect. *PloS one*.

Levere, M., Acharya, G., and Bharadwaj, P. (2016). The role of information and cash transfers on early childhood development: Evidence from Nepal. Working Paper No. 22640, National Bureau of Economic Research.

Macours, K., Premand, P., Schady, N., and Vakis, R. (2015). Experimental evidence from an early childhood parenting intervention in Nicaragua.

Macours, K., Schady, N., and Vakis, R. (2012). Cash transfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment. 4(2):247–273.

Macours, K. and Vakis, R. (2014). Changing households' investment behaviour through social interactions with local leaders: Evidence from a randomised transfer programme. *The Economic Journal*, 124(576):607–633.

Macours, K. and Vakis, R. (2017). Sustaining impacts when transfers end: Women leaders, aspirations, and investments in children. In *The Economics of Poverty Traps*, pages 325–355. University of Chicago Press.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *The Economic Journal*, 119(537):734–763.

Martinez, S., Naudeau, S., and Pereira, V. A. (2017). Preschool and child development under extreme poverty : evidence from a randomized experiment in rural Mozambique. Policy Research Working Paper No. 8290, The World Bank.

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., and Stampini, M. (2019). Long-term impacts of conditional cash transfers: Review of the evidence. *The World Bank Research Observer*, 34(1):119–159.

Mischel, W., Shoda, Y., and Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, 244(4907):933–938.

Nores, M., Belfield, C. R., Barnett, W. S., and Schweinhart, L. (2005). Updating the economic impacts of the High/Scope Perry preschool program. *Educational Evaluation and Policy Analysis*, 27(3):245–261.

Oster, E. and Thornton, R. (2012). Determinants of technology adoption: Peer effects in menstrual cup take-up. *Journal of the European Economic Association*, 10(6):1263–1293.

Ozler, B., Fernald, L. C., Kariger, P., McConnell, C., Neuman, M., and Fraga, E. (2018). Combining preschool teacher training with parenting education: A cluster-randomized controlled trial. *Journal of Development Economics*, 133:448=467.

Paxson, C. and Schady, N. (2007). Cognitive development among young children in Ecuador the roles of wealth, health, and parenting. *Journal of Human Resources*, 42(1):49–84.

Paxson, C. and Schady, N. (2010). Does money matter? The effects of cash transfers on child development in rural Ecuador. *Economic Development and Cultural Change*, 59(1):187–229.

Pop-Eleches, C., Thirumurthy, H., Habyarimana, J., Zivin, J., Goldstein, M., de Walque, D., and et al. (2011). Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS*, 25(6):825–834.

Powell, C., Baker-Henningham, H., Walker, S., Gernay, J., and Grantham-McGregor, S. (2004). Feasibility of integrating early stimulation into primary care for undernourished Jamaican children: cluster randomised controlled trial. *British Medical Journal*, 329(7457):89.

Premand, P. and Barry, O. (2020). Cash transfers, behavioral change promotion, and early childhood development: Experimental evidence from a government program in a low-income setting. Policy Research Working Paper No. 8670, Washington, D.C.: World Bank.

Schweinhart, L. J. (2005). *Lifetime effects: the High/Scope Perry Preschool study through age 40*. Number 14. High/Scope Foundation.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Valk, J.-H., Rashid, A. T., and Elder, L. (2010). Using mobile phones to improve educational outcomes: An analysis of evidence from Asia. *The International Review of Research in Open and Distributed Learning*, 11(1):117–140.

Walker, S. P., Chang, S. M., Powell, C. A., Simonoff, E., and Grantham-McGregor, S. M. (2007a). Early childhood stunting is associated with poor psychological functioning in late adolescence and effects are reduced by psychosocial stimulation. *The Journal of Nutrition*, 137(11):2464–2469.

Walker, S. P., Grantham-McGregor, S. M., Powell, C. A., and Chang, S. M. (2000). Effects of growth restriction in early childhood on growth, IQ, and cognition at age 11 to 12 years and the benefits of nutritional supplementation and psychosocial stimulation. *The Journal of Pediatrics*, 137(1):36–41.

Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., Carter, J. A., Group, I. C. D. S., et al. (2007b). Child development: Risk factors for adverse outcomes in developing countries. *The Lancet*, 369(9556):145–157.

Walsh, C. S., Power, T., Khatoon, M., Biswas, S. K., Paul, A. K., Sarkar, B. C., and Griffiths, M. (2013). The 'trainer in your pocket': Mobile phones within a teacher continuing professional development program in Bangladesh. *Professional Development in Education*, 39(2):186–200.

York, B. N., Loeb, S., and Doss, C. (2018). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. *Journal of Human Resources*.

7 Figures and Tables

FIGURE 1: Number of treated individuals participating in weekly quizzes over time

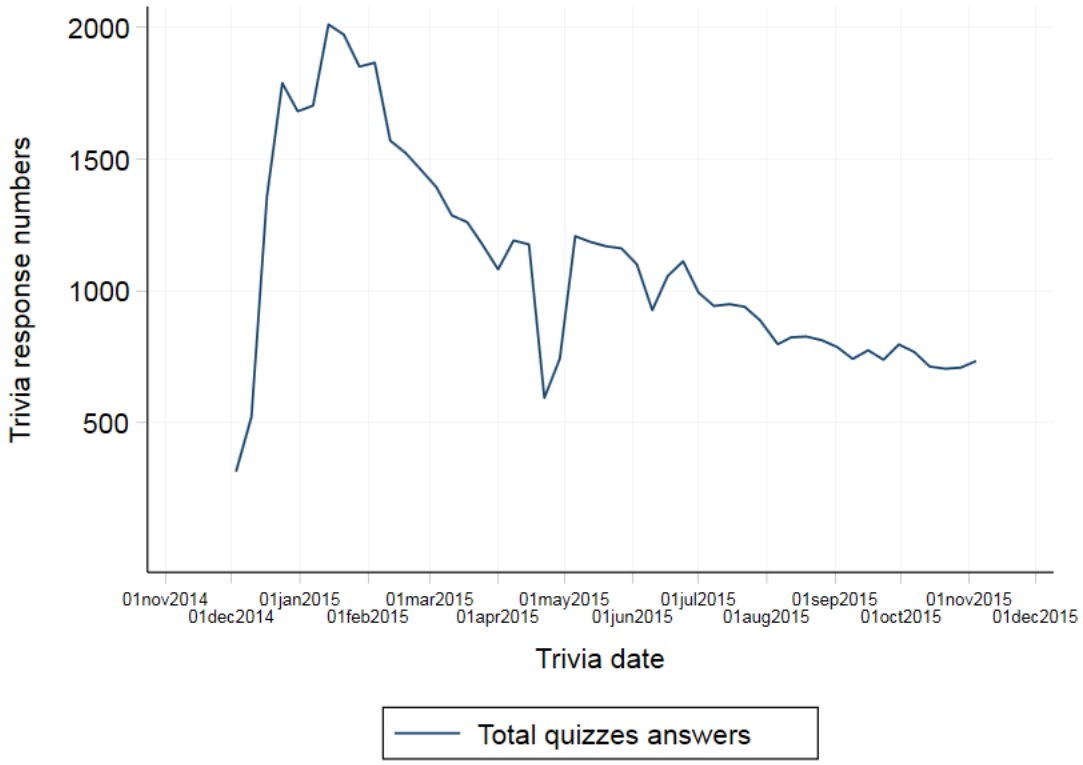


TABLE 1: Balance in baseline characteristics between treatment and control groups

	N	Control	Treatment	P-values with **stratification** (T - C)	P-Values **Unconditional** (T - C)
Child-specific characteristics					
Age at baseline	2,416	2.27	2.32	0.533	0.457
Father lived in household at baseline	2,416	0.76	0.75	0.770	0.535
Mother lived in household at baseline	2,416	0.96	0.96	0.916	0.969
Years education mother	2,414	3.26	3.27	0.468	0.889
Years education father	2,291	3.03	3.00	0.634	0.689
Health and Nutrition					
# Days in bed sick in last month	2,416	0.19	0.22	0.638	0.631
Received deworming in last 6 months	2,416	0.39	0.38	0.738	0.769
Received vitamins in last 6 months	2,416	0.39	0.38	0.786	0.801
# Days in the week the child had vegtables	2,415	1.71	1.88	0.068	0.074
# Days in the week the child had fruit	2,415	2.88	3.02	0.240	0.241
# Days in the week the child had meat	2,415	1.20	1.17	0.702	0.713
# Days in the week the child had eggs	2,414	2.69	2.64	0.689	0.653
# Days in the week the child had breast milk	2,414	2.46	2.45	0.936	0.972
# Days in the week the child had coffee	2,414	4.35	4.06	0.042	0.052
# Days in the week the child had soup	2,415	0.50	0.61	0.073	0.069
# Days read to the child in the last week	2,415	0.76	0.67	0.256	0.220
Someone read to the child in the last month	2,415	0.22	0.22	0.924	0.844
Household-level characteristics					
Male household head	2,512	0.86	0.86	0.443	0.827
Age household head	2,512	41.39	41.25	0.759	0.825
Literate household head	2,512	0.74	0.74	0.756	0.898
Number of men in the household	2,512	1.58	1.56	0.914	0.720
Number of women in the household	2,512	1.66	1.61	0.311	0.205
Number of boys (age 7-14) in the household	2,512	0.57	0.52	0.106	0.101
Number of girls (age 7-14) in the household	2,512	0.49	0.47	0.415	0.444
Number of young boys (age 0-6) in the hh	2,512	0.66	0.64	0.070	0.466
Number of young girls (age 0-6) in the hh	2,512	0.60	0.57	0.284	0.286
Number of rooms in house	2,511	1.91	1.84	0.139	0.111
Network					
Distance (in min) to closest cellphone signal	2,505	32.48	29.65	0.193	0.233
Distance (meters) to closest cellphone signal	2,343	2,116.53	1,979.93	0.832	0.478

Baseline survey characteristics of children and households who were visited at endline. The sample includes all baseline children with cognitive tests at follow-up and all households interviewed at follow-up. Children born after the baseline are not included. P-values are of the test of difference between control and treatment, after controlling for stratification variables (one-to-last column), or without any controls (i.e. unconditional, last column). P-values based on standard errors clustered by community.

TABLE 2: Baseline characteristics of villages with leaders assigned to text message intervention (treatment) or not (control)

	N	Control	Treatment	P-Values (T - C)
Village level characteristics				
Average number of leaders	92	4.59	4.23	0.557
Average number of households	92	34.81	30.22	0.277
Average age of the targeted children	92	2.36	2.41	0.509
Average number of households with fathers living in the household	92	24.59	22.11	0.401
Average number of households with mothers living in the household	92	32.30	27.94	0.263
Average number of households in the government ECD program	92	4.78	4.69	0.952
Average number of households with access to electricity	92	19.41	15.69	0.457
Avg years of education fathers	92	2.78	2.64	0.331
Avg years of education mothers	92	3.14	3.01	0.352

The treatment group is composed of villages where opinion leader households were selected to receive text messages, the control group is composed of villages where opinion leader households were not selected to receive text messages. Five villages without opinion leaders in the sample are not included.

TABLE 3: Impact of text messages on Early Childhood Development outcomes

	(1) ECD full sample	(2) ECD young	(3) ECD Old
Panel A: Intention to treat			
ITT	0.00	-0.01	0.00
	(0.03)	(0.06)	(0.04)
Panel B: Heterogeneity by caregiver education			
ITT	-0.12**	-0.24*	-0.07
	(0.06)	(0.13)	(0.07)
ITT X Medium educ (4-6 years of education)	0.17**	0.18	0.13
	(0.08)	(0.16)	(0.09)
ITT X High educ (more than primary education)	0.18**	0.41***	0.04
	(0.08)	(0.15)	(0.09)
Observations	2485	774	1711
P-values joint significance test			
ITT + ITT X Medium educ = 0	0.35	0.48	0.22
ITT + ITT X High educ = 0	0.27	0.07	0.74

Note: All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old, and the first principal component of all 7 tests for children 36-84 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Standard errors reported in parentheses.* p<0.1, ** p<0.05, *** p<0.01

TABLE 4: Impact of text messages on intermediate outcomes

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
Panel A. ITT							
ITT	0,06	0,16***	0,07*	0,08*	0,06	0,11**	0,09**
	(0,04)	(0,04)	(0,04)	(0,04)	(0,04)	(0,04)	(0,04)
Panel B. By treatment variation (thematic focus)							
Nutrition & Stimulation	0,01	0,20***	0,11**	0,05	0,04	0,16***	0,08
	(0,05)	(0,05)	(0,05)	(0,05)	(0,05)	(0,06)	(0,05)
Nutrition	0,10**	0,05	0,00	0,13**	0,07	-0,02	0,09*
	(0,05)	(0,05)	(0,05)	(0,05)	(0,05)	(0,06)	(0,05)
Stimulation	0,08	0,23***	0,08	0,03	0,06	0,19***	0,10*
	(0,05)	(0,05)	(0,05)	(0,06)	(0,05)	(0,06)	(0,05)
Observations	2500	2501	2512	2495	2501	2396	2506
P-values Statistical Significance Test							
Nutrition & Stimulation = Nutrition	0.14	0.00	0.02	0.15	0.52	0.00	0.94
Nutrition & Stimulation = Stimulation	0.25	0.58	0.10	0.81	0.70	0.56	0.66
Stimulation = Nutrition	0.76	0.00	0.57	0.10	0.80	0.00	0.71

Note: All outcome variables are standardized using the mean and standard deviation of the control households. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. Impacts on each of the individual items used to construct the indices are reported in tables A10, A11, A12. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Standard errors reported in parentheses.* p<0.1, ** p<0.05, *** p<0.01

TABLE 5: Impact of leaders' exposure on Early Childhood Development outcomes

	(1) ECD full sample	(2) ECD young	(3) ECD Old
Panel A ITT of leaders' exposure			
ITT Leader	-0.11***	-0.02	-0.14***
	(0.04)	(0.05)	(0.05)
Panel B Heterogeneity by caregiver education			
ITT Leader	-0.20***	-0.22*	-0.18***
	(0.06)	(0.12)	(0.07)
ITT Leader X Medium educ (4-6 years of education)	0.17**	0.35**	0.09
	(0.07)	(0.16)	(0.08)
ITT Leader X High educ (more than primary education)	0.08	0.17	0.01
	(0.07)	(0.14)	(0.10)
Observations	2051	624	1427
P-values joint significance test			
ITT Leader + ITT Leader X Medium educ = 0	0,58	0,20	0,13
ITT Leader + ITT Leader X High educ = 0	0,03	0,49	0,05

Note: Sample includes only non-leader households in 92 villages with opinion leaders. Five villages without leaders were excluded from the estimation. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, child age and gender, and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE 6: Impact of leaders' exposure by distance to the closest leader's dwelling

	(1) ECD full sample	(2) ECD young	(3) ECD Old
Panel A: Leader treatment interacted with distance from closest leader			
ITT Leader	-0.168***	-0.126*	-0.195***
	(0.05)	(0.07)	(0.07)
Distance from closest leader	-0.009***	-0.017***	-0.006
	(0.00)	(0.00)	(0.00)
ITT leader X distance from closest leader	0.012***	0.022***	0.010*
	(0.00)	(0.01)	(0.01)
Panel B: Leader treatment interacted with household treatment status			
ITT	-0.046	0.036	-0.112*
	(0.06)	(0.13)	(0.06)
ITT Leader	-0.151**	0.029	-0.251***
	(0.06)	(0.13)	(0.07)
ITT X ITT Leader	0.060	-0.071	0.150*
	(0.07)	(0.15)	(0.08)
Observations	2051	624	1427
P-vales joint significance test			
ITT Leader + ITT Leader X ITT	0,04	0,51	0,07

Note: Distance from leader is measured in terms of the number of houses between the household and the closest household with an opinion leader. Sample includes only non-leader households in 92 villages with opinion leaders. Five villages without leaders were excluded from the estimation. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender, and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE 7: Impact of leaders' exposure on intermediate outcomes

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
Panel A: ITT of leaders' exposure to text messages							
ITT Leader	-0,14*	0,02	-0,00	-0,16**	-0,23**	-0,00	-0,11*
	(0,08)	(0,06)	(0,04)	(0,06)	(0,09)	(0,06)	(0,06)
Observations	2062	2062	2073	2059	2062	1970	2068
Panel B: Heterogeneity by caregiver education							
ITT Leader	-0,01	0,10	-0,01	-0,23***	-0,18*	0,00	-0,08
	(0,12)	(0,11)	(0,07)	(0,09)	(0,09)	(0,10)	(0,09)
ITT Leader X Medium educ (4-6 years of education)	-0,04	-0,04	-0,01	0,08	0,03	0,04	-0,05
	(0,10)	(0,12)	(0,08)	(0,09)	(0,09)	(0,11)	(0,10)
ITT Leader X High educ (more than primary education)	-0,34***	-0,21	0,04	0,10	-0,20	-0,06	-0,04
	(0,13)	(0,13)	(0,10)	(0,10)	(0,12)	(0,13)	(0,10)
Observations	2062	2062	2073	2059	2062	1970	2068
Joint significance test							
ITT Leader + ITT Leader X Medium Educ	0,58	0,46	0,66	0,08	0,12	0,61	0,16
ITT Leader + ITT Leader X High Educ	0,00	0,12	0,67	0,16	0,01	0,60	0,13

Note: Sample includes only non-leader households in 92 villages with opinion leaders. Five villages without leaders were excluded from the estimation. All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. Impacts on each of the individual items used to construct the indices are reported in Tables A17, A18, A19. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE 8: Impact of leaders' exposure on ECD outcomes, by leaders' prior beliefs

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT Leader	-0.08*	-0.09	-0.08
	(0.05)	(0.06)	(0.05)
Agreeing leader	0.07	-0.10	0.15*
	(0.07)	(0.09)	(0.08)
ITT leader X Agreeing leader	-0.06	0.16	-0.16
	(0.08)	(0.11)	(0.10)
Observations	2051	624	1427
ITT Leader + ITT leader X Agreeing leader = 0	0.04	0.44	0.01

Note: Agreeing leader is a binary variable indicating that the closest leader agreed with at least 4 out of 5 program messages on parenting practices at baseline. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old, and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE 9: Impact of leaders' exposure on intermediary outcome,s by leaders' prior beliefs

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
ITT Leader	-0.12	0.06	-0.01	-0.25***	-0.25***	0.01	-0.10
	(0.09)	(0.06)	(0.04)	(0.07)	(0.09)	(0.07)	(0.08)
Agreeing leader	0.09	0.11	0.07	-0.11	0.06	0.01	0.06
	(0.09)	(0.09)	(0.06)	(0.10)	(0.10)	(0.08)	(0.10)
ITT leader X Agreeing leader	-0.04	-0.12	0.01	0.22*	0.05	-0.03	-0.02
	(0.11)	(0.11)	(0.07)	(0.12)	(0.12)	(0.10)	(0.12)
Observations	2062	2062	2073	2059	2062	1970	2068
ITT Leader + ITT leader X Agreeing leader = 0	0.12	0.58	0.96	0.77	0.11	0.82	0.19

Note: Agreeing leader is a binary variable indicating that the closest leader agreed with at least 4 out of 5 program messages on parenting practices at baseline. All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE 10: Impact of leaders' exposure on ECD outcomes, by closest leader's commitment

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT Leader	-0,12**	-0,01	-0,17***
	(0,05)	(0,07)	(0,06)
Committed leader	-0,01	0,05	-0,04
	(0,05)	(0,13)	(0,05)
ITT Leader X Committed leader	0,08	-0,03	0,12
	(0,09)	(0,17)	(0,09)
Observations	2051	624	1427
ITT Leader + ITT Leader X Committed leader	0,56	0,74	0,59

Note: Committed leader is a leader predicted to be in the top quintile of leaders with highest participation in the quizzes. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 and 35 month. In column 3, the dependent variables is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE 11: Impact of leaders' exposure on intermediary outcomes, by closest leader's commitment

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
ITT Leader	-0.16*	-0.01	-0.02	-0.13*	-0.25**	-0.05	-0.13*
	(0.08)	(0.06)	(0.04)	(0.07)	(0.10)	(0.06)	(0.07)
Committed leader	-0.15	-0.16	-0.06	0.19**	-0.10	-0.25***	-0.15
	(0.12)	(0.11)	(0.07)	(0.08)	(0.10)	(0.08)	(0.13)
ITT leader X Committed leader	0.09	0.08	0.10	-0.11	0.11	0.23*	0.06
	(0.17)	(0.14)	(0.09)	(0.11)	(0.14)	(0.12)	(0.15)
Observations	2062	2062	2073	2059	2062	1970	2068
ITT Leader + ITT Leader X Committed leader	0.67	0.61	0.38	0.02	0.26	0.14	0.65

Note: Committed leader is a leader predicted to be in the top quintile of leaders with highest participation in the quizzes. All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE 12: Impacts on social interactions about ECD with targeted leaders

	(1) # leaders	(2) Health promotor	(3) Pre-sch Teacher	(4) Primary Teacher	(5) Other Educator	(6) Local leader	(7) Other com. leader
ITT	-0.01	0.02	-0.04	-0.01	-0.01	0.01	0.00
	(0.05)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
ITT Leader	-0.18**	-0.03	-0.03	-0.07***	-0.02**	-0.02	-0.02**
	(0.07)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Observations	2073	2073	2073	2073	2073	2073	2073
Mean Control	0.89	0.13	0.33	0.24	0.05	0.10	0.03

Note: The dependent variables are the social interactions between the household and targeted opinion leaders during the previous week. Column 1 is an index accounting for the social interaction between the household and health promoters, preschool-teachers, primary school teachers, other teachers, the local leader or other community leader. The dependent variables in the remaining columns are binary and indicate whether the household had at least one interactions with each type of leader during the previous week. All regressions include controls for the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects, as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

A1 Online Appendix

FIGURE A1: Number of treated individuals participating in weekly quizzes over time, by content of messages

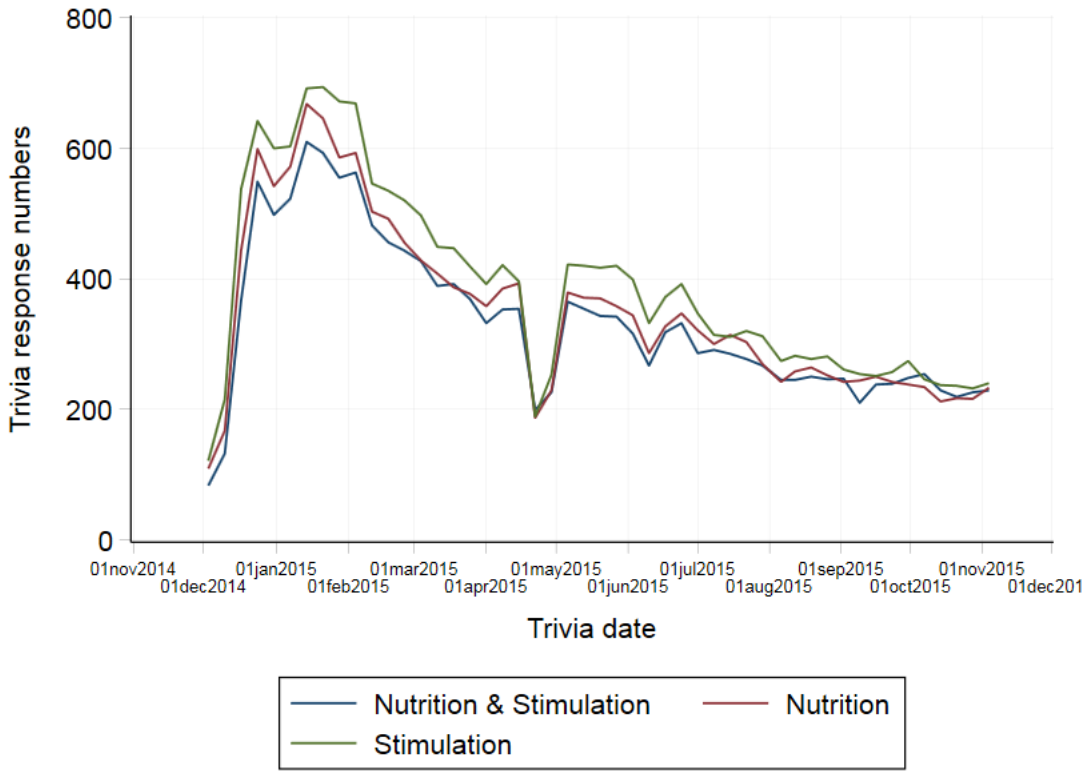
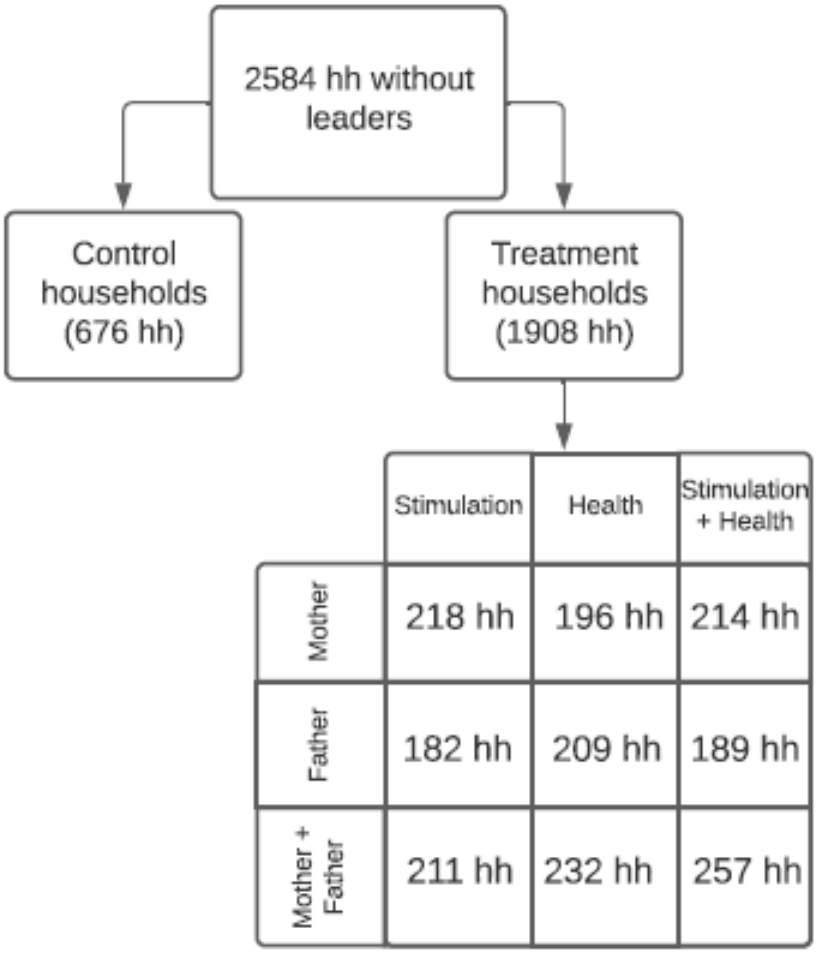
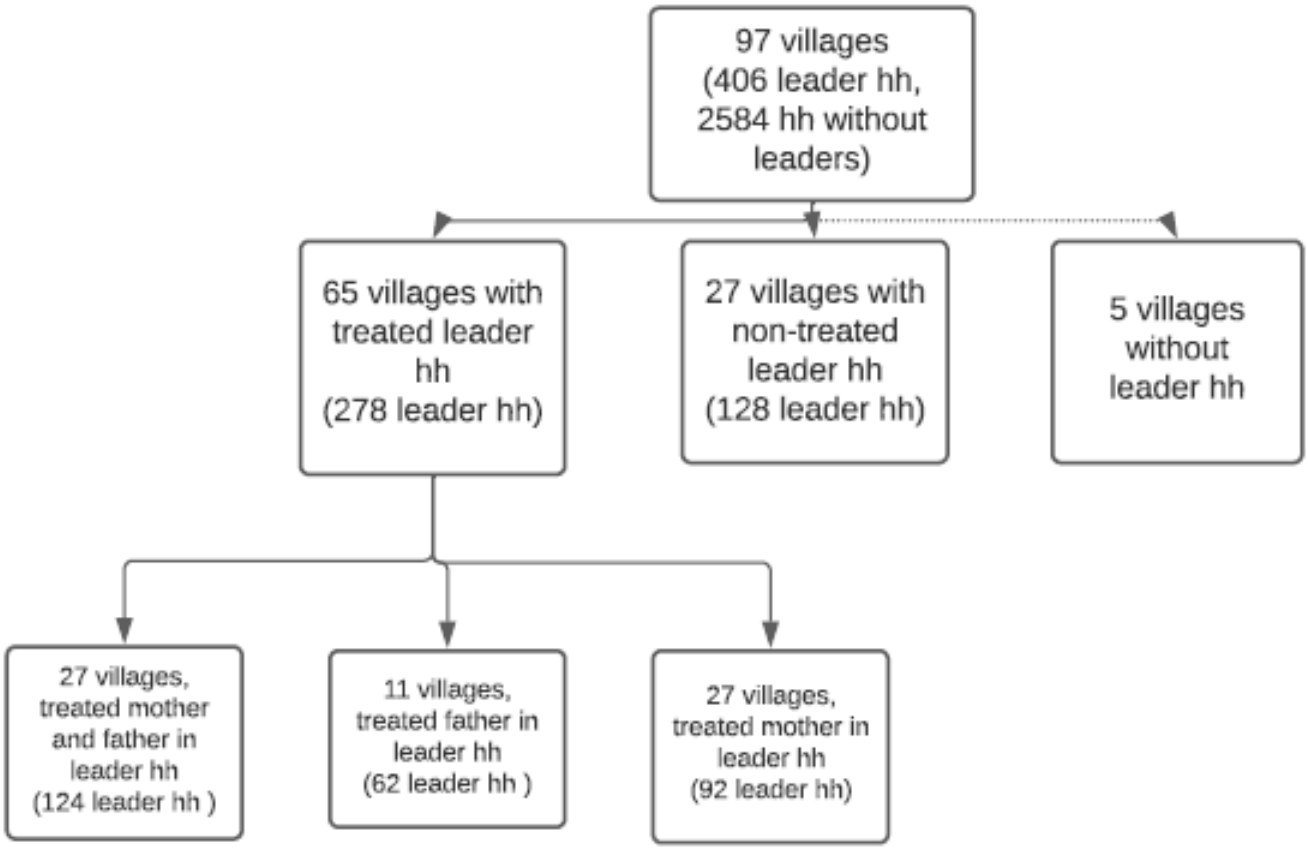


FIGURE A2: Household-level randomization



Note: the variation in the number of households with each treatment variation is due to stratification

FIGURE A3: Village-level randomization



A1.1 Supplementary tables

TABLE A1: Example of text messages, by content, for children 7-12 months old

Content	Message
Nutrition and Health	Take NAME to her check-ups every two months on time
	After 6 months, start feeding NAME little by little
	Vaccine NAME against dangerous diseases so that she grows healthy
	Give NAME mashed vegetables after 6 months
	Bathe NAME daily. Keep him / her clean so he/she doesn't get sick
	Breastfeed NAME in addition to giving her other foods until she is at least 2 years old
	Take care of NAME if she has a cough. A cough that lasts is dangerous
	Give NAME mashed vegetables (pipian, chayote, carrot, squash,...)
	Teach NAME to brush her teeth so they are always clean and healthy
	Give NAME vegetables, they have a lot of vitamins and are nutritious
	Put slippers or shoes on NAME so worms don't bother her
	Give NAME fruits in small pieces to help him grow healthy
	Wash NAME's hands often so she doesn't get sick
	Protect NAME from malaria and dengue by sleeping under a mosquito net
	Give NAME a well cut or cooked tomato, it is good for her to grow well
	Let your house be fumigated so there are no mosquitoes
	After 6 months, give NAME mashed fruits, papaya and mango
	Don't cook with wood near NAME. The smoke makes the lungs sick
	Do not give NAME candy and chips, they are junk food that does not feed him
	Keep NAME clean so she doesn't get sick
	Wash NAME's clothes well so that he is clean and healthy
	Don't give NAME soda or packaged juices, they are not good for her
	After 6 months, you can already give NAME beans
	Wash NAME's hands several times a day
	Give NAME your milk and about 12 tablespoons of mashed food 3 to 4 times a day
	Don't smoke near NAME, it hurts her
	Feed NAME vegetables and leafy greens: they are very nutritious
	Protect NAME from flies, they transmit diseases
	Feed NAME cabbage, spinach and lettuce, they have iron and they are good
	After 6 months, give NAME other mashed foods
Stimulation and Home Environment	Speak with affection to NAME
	Give NAME hugs, children need a lot of affection
	Put NAME with his hands together and play with him to try to separate them
	Do not mistreat NAME so that he can love you and grow up healthy and happy
	Read stories to NAME
	Don't yell at NAME, the screams hurt
	Make simple toys for NAME to push or drag using what you have at home
	After 11 months, help NAME to start walking
	Do not compare NAME with other children. All children are different
	Take NAME for a walk and lovingly show him different things
	Do not say ugly names or curse NAME. That hurts more than a blow
	Tell stories to NAME
	While you work around the house put NAME in a safe place
	Make toys for NAME. A tin with stones inside is good for playing
	Make NAME her little corner in the house so she has her toys in order
	Let NAME touch, hear and observe different things and situations
	Curiosity is good for learning and developing
	Keep a place in the house clear of furniture so that NAME can crawl and walk
	Make NAME listen to different sounds and tell him where each sound comes from
	Smile at NAME, this way she will be happier
	Caress NAME, touch her arms, head and legs gently, massage her back so that she feels loved
	Give NAME tenderness so that he grows up feeling safe
	Take NAME to see the sun when it rises or sets. Explain what the sun is for
	Please bear with NAME when she cries. Remember that he can't say what he wants
	Make NAME touch different things to make him feel how they are
	Explain to NAME what the things he touches are for. It educates him
	Don't let animals stay where NAME plays so she doesn't get sick
	Keep the places where NAME spends the most time clean
	Make simple toys for NAME with whatever you have on hand

Note: Translated from Spanish.

TABLE A2: Principal component analysis: ECD (cognitive) outcomes

	(1) Variance	(2) Proportion	(3) Factor loadings
Denver subcomponents (used for children 1-3 years old)	1.390	0.3476	
Personal - Social			0.385
Language			0.651
Fine Motor			0.675
Gross Motor			0.601
All tests together (used for children 3-7 years old)	2.268	0.324	
Denver Personal - Social			0.100
Denver Language			0.744
Denver Fine Motor			0.595
Denver Gross Motor			0.486
Vocabulary (TVIP)			0.754
Memory			0.735
Self-control			0.100

Note: Column 1 reports the eigenvalue of the first principal component. Column 2 reports the proportion of variance accounted for by the first principal component. Column 3 reports the principal component factor loadings, which show the correlation between the variables and the first principal component.

TABLE A3: Principal component analysis: socio-emotional outcomes

	(1) Variance	(2) Proportion	(3) Factor loadings
Behavioral (used for children 1-3 years old)	1.320	0.440	
Control			0.504
Affect			0.761
Ability			0.697
Strengths and Difficulties (used for children 3-7 years old)	1.718	0.245	
Emotional			0.687
Conduct			0.783
Hyperactivity			0.124
Peer			0.733
Pro-social			-0.075
Control			0.262
Laugh			-0.059

Note: Column 1 reports the eigenvalue of the first principal component. Column 2 reports the proportion of variance accounted for by the first principal component. Column 3 reports the principal component factor loadings, which show the correlation between the variables and the first principal component.

TABLE A4: ITT effects on attrition

	(1) Child	(2) Household	(3) Mother	(4) Father
Panel A				
ITT	0.022*	0.015	0.010	0.016
	(0.013)	(0.012)	(0.014)	(0.015)
Observations	2803	2803	2803	2803
Mean control	0.089	0.071	0.115	0.199
Panel B				
Nutrition & Stimulation	0.018	0.012	0.000	0.006
	(0.017)	(0.015)	(0.018)	(0.019)
Nutrition	0.039**	0.025*	0.029	0.035*
	(0.017)	(0.015)	(0.018)	(0.019)
Stimulation	0.017	0.009	0.015	0.025
	(0.017)	(0.015)	(0.018)	(0.019)
Observations	2803	2803	2803	2803
Mean control	0.089	0.071	0.115	0.199
P-values Statistical Significance Test				
Nutrition & Stimulation = Nutrition	0.261	0.389	0.136	0.144
Nutrition & Stimulation = Stimulation	0.952	0.890	0.460	0.362
Stimulation = Nutrition	0.245	0.325	0.474	0.610
Panel C				
ITT Leader	-0.004	-0.001	-0.010	-0.013
	(0.017)	(0.015)	(0.021)	(0.016)
Observations	2332	2332	2332	2332
Mean Control	0.105	0.082	0.125	0.219

Note: The dependent variable measures the probability of attrition for children's test outcomes (column 1), and information obtained from the main caregiver in the household (2), mothers (3) and fathers (4). Coefficients in panel A and B are estimated on the full sample of children less than 7 years old at baseline using equation 1 and equation 2 (including controls for the stratification variables, child age and gender, enumerator fixed effects, as well as village fixed effects). Coefficients in panel C are estimated on children less than 7 years old from non-leader households only using equation 3 (including controls for the household-level treatment, the average level of education of the leaders, as well as the level of education of the main caregiver, whether there is a male caregiver in the household, child age and gender, and test-administrator fixed effects). The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE A5: Balancing test on attrited observations

	N	Control	Treatment	P-values with stratification (T - C) with controls	P-Values Unconditional (T - C)
Children without test-results at follow-up					
Age at the base line	274	2.23	2.09	0.228	0.464
# Days in bed over last month	274	0.17	0.13	0.430	0.678
Received deworming drugs over last 6 months	274	0.45	0.38	0.545	0.303
Received vitamins over last 6 months	274	0.52	0.38	0.033	0.056
# Days in the week the child had vegtables	274	1.81	1.60	0.317	0.483
# Days in the week the child had fruit	274	2.77	2.69	0.962	0.842
# Days in the week the child had meat	274	1.36	1.31	0.738	0.832
# Days in the week the child had eggs	274	2.42	2.66	0.632	0.507
# Days in the week the child was breast fed	274	2.53	2.43	0.475	0.838
# Days in the week the child had coffee	274	3.52	4.00	0.329	0.313
# Days in the week the child had soup	274	0.77	0.47	0.025	0.057
# Days read to the child in the last week	274	0.88	0.59	0.248	0.190
Someone read to the child in the last month	274	0.22	0.21	0.877	0.875
Attrited households					
Male household head	223	0.93	0.88	0.949	0.363
Age household head	223	41.33	41.18	0.980	0.953
Literacy household head	223	0.78	0.76	0.714	0.828
Number of men in the household	223	1.56	1.59	0.657	0.826
Number of women in the household	223	1.63	1.64	0.894	0.947
Number of boys (age 7-14) in the household	223	0.31	0.46	0.283	0.159
Number of girls (age 7-14) in the household	223	0.43	0.43	0.693	0.999
Number of young boys (age 0-6) in the hh	223	0.59	0.62	0.662	0.797
Number of young girls (age 0-6) in the hh	223	0.50	0.59	0.886	0.360
Number of rooms in the house	223	2.07	1.69	0.086	0.010
Attrited mothers					
Mother lived in household at baseline	340	0.02	0.03	0.866	0.668
Years education mother	287	1.21	1.09	0.860	0.554
Male household head	340	0.89	0.84	0.602	0.268
Age household head	340	43.73	42.31	0.355	0.477
Literacy household head	340	0.75	0.76	0.282	0.882
Number of males in the household	340	1.64	1.57	0.493	0.628
Number of women in the household	340	1.76	1.75	0.445	0.906
Number of boys (7-14) in the household	340	0.44	0.43	0.733	0.901
Number of girls (7-14) in the household	340	0.44	0.39	0.239	0.519
Number of young boys (0-6) in the hh	340	0.73	0.66	0.697	0.386
Number of young girls (0-6) in the hh	340	0.49	0.56	0.884	0.334
Number of rooms in house	340	2.11	1.79	0.048	0.012
Attrited fathers					
Father lived in household at baseline	619	0.01	0.01	0.849	0.977
Years education father	481	0.97	1.09	0.791	0.489
Male household head	620	0.59	0.63	0.184	0.455
Age household head	620	43.13	43.93	0.682	0.594
Literacy household head	620	0.74	0.74	0.246	0.910
Number of men in the household	620	1.39	1.40	0.701	0.884
Number of women in the household	620	1.82	1.88	0.766	0.559
Number of boys (age 7-14) in the household	620	0.52	0.49	0.399	0.625
Number of girls (age 7-14) in the household	620	0.37	0.46	0.448	0.141
Number of young boys (age 0-6) in the hh	620	0.73	0.68	0.120	0.398
Number of young girls (age 0-6) in the hh	620	0.52	0.56	0.815	0.473
Number of rooms in house	620	2.13	1.81	0.017	0.001

All data from 2014 baseline survey. P-values based on standard errors clustered by community. The number of observations (N) indicating child-level attrition is the difference between the number of baseline households and the number of children for whom the Denver test was completed; N for household-level attrition is the number of baseline households for whom no follow up household survey was collected; and N for mother(father)-level is the number of baseline households for whom the mother(father) could not be interviewed at follow-up.

TABLE A6: Impact of text messages on individual ECD tests

	ITT	**S.e**	**P-Value**	**Obs**
	(1)	(2)	(3)	(4)
Younger children (12 - 36 months)				
ECD index (first principal component)	-0.01	(0.06)	0.87	774
Denver Personal-social	0.02	(0.08)	0.82	790
Denver Language	0.04	(0.05)	0.39	784
Denver Fine Motor	0.01	(0.07)	0.87	781
Denver Gross Motor	-0.13	(0.10)	0.17	779
Older children (37 - 84 months)				
ECD index (first principal component)	0.00	(0.04)	0.95	1,711
Denver Personal-social	0.01	(0.04)	0.81	1,757
Denver Language	0.06	(0.04)	0.81	1,753
Denver Fine Motor	0.04	(0.05)	0.37	1,746
Denver Gross Motor	-0.01	(0.05)	0.80	1,742
Memory	-0.05	(0.04)	0.21	1,927
TVIP (Vocabulary)	-0.03	(0.04)	0.41	1,757
Self control	0.00	(0.05)	0.93	1,713

Each row corresponds to a separate estimation (equation 1). All outcome variables are standardized using the mean and standard deviation of control households. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variables for a set of households surveyed in 2016, as well as community fixed effects. Standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE A7: Impacts of text messages on ECD Outcomes, by treatment variations

	(1) ECD full sample	(2) ECD young	(3) ECD Old
Panel A: By treatment variation (thematic focus)			
Nutrition & Stimulation	-0,00	-0,05	0,02
	(0,04)	(0,07)	(0,05)
Nutrition	-0,02	0,02	-0,05
	(0,04)	(0,08)	(0,05)
Stimulation	0,02	0,00	0,02
	(0,04)	(0,08)	(0,05)
Observations	2485	774	1711
Panel B: By treatment variation (targeted caregiver)			
Mother	0,03	0,05	0,02
	(0,04)	(0,08)	(0,05)
Father	-0,01	-0,01	-0,01
	(0,04)	(0,08)	(0,05)
Mother and Father	-0,02	-0,06	-0,02
	(0,04)	(0,07)	(0,05)
Observations	2485	774	1711

Note: All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months and the first principal component of all 7 tests for children 36-84 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effect. Standard errors reported in parentheses.* p<0.1, ** p<0.05, *** p<0.01

TABLE A8: Impacts of text messages on socio-emotional outcomes for children aged 3-7 years

	(1) Behavioral Index	(2) Emotional	(3) Conduct	(4) Hyperactive	(5) Peer	(6) Pro Social	(7) Control	(8) Laugh and Smile
Panel A: Equation 1								
ITT	- 0.02	0.08	-0.10*	0.10*	0.00	-0.01	-0.03	0.02
	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)
Observations	1723	1723	1723	1723	1723	1723	1723	1723
Panel B: Equation 2								
Nutrition & Stimulation	0.04	0.16**	-0.03	0.13*	-0.02	0.01	0.00	0.06
	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)	(0.07)
Nutrition	-0.03	0.04	-0.11	0.05	0.04	-0.03	-0.09	0.00
	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
Stimulation	-0.03	0.07	-0.16**	0.09	0.02	0.00	-0.00	-0.00
	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
Observations	1723	1723	1723	1723	1723	1723	1723	1723
Panel C: Equation 3								
ITT Leader	-0.05	-0.00	-0.05	0.02	-0.06	-0.00	-0.03	-0.13**
	(0.06)	(0.05)	(0.07)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)
Observations	1438	1438	1438	1438	1438	1438	1438	1438

Note: The dependent variable in the first column is the first principal component of the 5 subscales of the "Strengths and difficulties" questionnaire, and modules measuring inhibitory control scale and positive demeanor (Laugh and Smile). Controls in panel A and B as in Table 3, controls in panel C as in Table 4. The standard errors (in parentheses).* p<0.1, ** p<0.05, *** p<0.01

TABLE A9: Impacts of text messages on socio-emotional outcomes for children aged 1-3 years

	(1) Behavioral index	(2) Control	(3) Affect	(4) Ability
Panel A: Equation 1				
ITT	0.02	-0.14	0.04	0.09
	(0.08)	(0.09)	(0.09)	(0.09)
Observations	717	717	717	717
Panel B: Equation 2				
Nutrition & Stimulation	- 0.01	-0.16	0.01	0.10
	(0.11)	(0.11)	(0.11)	(0.11)
Nutrition	0.07	-0.27**	0.11	0.22**
	(0.11)	(0.11)	(0.11)	(0.11)
Stimulation	0.02	0.05	-0.00	0.01
	(0.12)	(0.11)	(0.11)	(0.11)
Observations	717	717	717	717
Panel C: Equation 3				
ITT Leader	0.02	-0.01	0.04	0.00
	(0.10)	(0.09)	(0.09)	(0.11)
Observations	577	577	577	577

Note: The dependent variable in the first column is the first principal component of 3 sub-scales of the early childhood behavior questionnaire: effortful control (column 2), negative affectivity (column 3) and sociability (or extraversion, column 4). Controls in panel A and B as in Table 3, controls in panel C as in Table 4. Standard errors (in parentheses).* p<0.1, ** p<0.05, *** p<0.01

TABLE A10: Impact of text messages on individual intermediary outcomes: nutrition and stimulation

	ITT	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Nutrition					
Nutrition index	0,06	(0,04)	0,12	0,00	2.500
Beverages and soups					
Soy Milk	0,06	(0,06)	0,33	0,37	2.501
Juice	0,16	(0,11)	0,15	3,72	2.501
Soup	0,00	(0,06)	0,96	0,71	2.501
Vegetables					
Salad	-0,13	(0,09)	0,17	0,34	2.501
Potatoes	0,10	(0,07)	0,18	1,08	2.501
Tomatoes	0,06	(0,12)	0,58	2,97	2.501
Onion	0,22	(0,13)	0,10	1,98	2.501
Other vegetables	-0,02	(0,09)	0,85	1,69	2.501
Staples					
Rice	0,03	(0,08)	0,73	5,85	2.500
Beans	-0,06	(0,09)	0,49	6,09	2.501
Bread	-0,02	(0,12)	0,85	4,12	2.501
Tortilla	0,11	(0,10)	0,28	5,82	2.499
Cookies	0,13	(0,09)	0,16	1,28	2.501
Proteins					
Protein index	0,06	(0,04)	0,11	0,00	2.501
Milk	-0,04	(0,11)	0,74	1,76	2.501
Eggs	0,30	(0,09)	0,00	1,72	2.501
Cheese	-0,00	(0,10)	0,98	2,17	2.500
Meat	-0,00	(0,06)	0,98	1,22	2.501
Breast milk	0,22	(0,25)	0,38	3,15	802
Stimulation					
Stimulation index	0,16	(0,04)	0,00	0,00	2.501
Buy toys for the child	0,03	(0,01)	0,01	0,86	2.501
Tell tales to the child	0,04	(0,02)	0,02	0,79	2.501
Read books to child	0,04	(0,02)	0,05	0,27	2.501
Have pen and paper	0,02	(0,02)	0,12	0,84	2.501

Each row corresponds to a separate estimation (equation 1) and shows the ITT estimates on aggregate indices or individual questions measuring investments in nutrition or stimulation as reported by the caregiver. Questions on food items measure how many days in the last 7 days the child was given the specific item. The nutrition index was calculated by summing the days over all food items. The protein index was calculated by summing the days over all protein items. The individual items about stimulation are answers to yes/no questions, and the index is obtained by summing over yes answers. The nutrition, protein, and stimulation indices are standardized using the mean and standard deviation for control households. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

TABLE A11: Impact of text messages on individual intermediary outcomes: preventive health

	ITT	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Health					
Health index	0,07	(0,04)	0,07	0,00	2.512
Insist on brushing teeth	0,01	(0,01)	0,41	0,02	2.512
Ensure a healthy diet	0,03	(0,01)	0,02	0,06	2.512
Keep child away from stove	-0,00	(0,00)	0,93	0,00	2.512
Ensure child's good hygiene	-0,01	(0,02)	0,46	0,21	2.512
Wash child's clothes properly	0,01	(0,02)	0,60	0,12	2.512
Avoid giving coffee	-0,00	(0,00)	0,32	0,00	2.512
Make child wash hands	0,00	(0,02)	0,86	0,27	2.512
Smoke at a distance from child	-0,00	(0,00)	0,74	0,00	2.512
Do not give child unhealthy snacks	0,00	(0,01)	0,49	0,01	2.512
Cover fresh food	0,00	(0,02)	0,78	0,13	2.512
Teach child to chew slowly	-0,00	(0,00)	0,61	0,00	2.512
Prevent child from catching the flu	0,00	(0,01)	0,80	0,06	2.512
Boil or chlorinate the water	0,01	(0,01)	0,13	0,02	2.512
Ensure that child vaccines are up-to-date	0,02	(0,01)	0,23	0,07	2.512
Use a mosquito net	-0,00	(0,01)	0,71	0,01	2.512
Insist child wears shoes	0,01	(0,01)	0,56	0,03	2.512
Make sure child eats regularly	0,00	(0,00)	0,25	0,00	2.512
Keep child away from dangerous products	0,00	(0,00)	0,75	0,00	2.512
Show child affection	-0,01	(0,00)	0,09	0,01	2.512
Other	-0,00	(0,02)	0,97	0,17	2.512

Every row corresponds to a separate estimation (equation 1) and shows the ITT estimates on the aggregate health index or on individual questions measuring whether the caregiver reported a given practice when asked how to avoid their young child getting sick. The health index is calculated based on the number of different preventive health practices. The index is standardized using the mean and standard deviation of the control households. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Standard errors reported in parentheses.* p<0.1, ** p<0.05, *** p<0.01.

TABLE A12: ITT on individual intermediary outcomes: micronutrients, ECD attitudes and observed hygiene

	ITT	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Micronutrients					
Micronutrients index	0.08	(0.04)	0.06	-0.00	2,495
Received Vitamin A	0.02	(0.02)	0.21	0.73	2,495
Received deworming medecine	0.02	(0.02)	0.37	0.76	2,495
Received iron supplementation	0.04	(0.02)	0.07	0.49	2,495
ECD Attitudes of main caregiver					
Attitudes index	0.11	(0.04)	0.01	0.00	2,396
Should always answer child's questions	0.02	(0.02)	0.28	0.87	2,396
Boys can play with dolls	0.07	(0.02)	0.00	0.36	2,396
Girls can play with cars	0.04	(0.02)	0.04	0.51	2,396
Early language develops from stimulation	-0.00	(0.02)	0.87	0.67	2,396
Brain develops from very early on	0.01	(0.02)	0.51	0.86	2,396
Hygiene (observed by test administrator)					
Hygiene index	0.09	(0.04)	0.03	0.00	2,506
Child has clean face	0.03	(0.02)	0.09	0.67	2,506
Child has clean hair	-0.00	(0.02)	0.83	0.80	2,506
Child has clean hands	0.06	(0.02)	0.00	0.56	2,506
Child has clean clothes	0.07	(0.02)	0.00	0.56	2,506
Child is wearing shoes	-0.04	(0.02)	0.05	0.33	2,506
Child does not cough	-0.01	(0.02)	0.40	0.85	2,506
Child has clean nose	0.00	(0.02)	0.99	0.84	2,506
Child does not have skin problems	0.00	(0.01)	0.83	0.93	2,505

Every row corresponds to a separate estimation (equation 1). The vitamin index is based on three variables: the child has received vitamin A, ferrous sulfate and deworming medicine during the last six months (yes/no questions). The hygiene index captures the hygienic condition of the child observed by the test administrator. The Attitudes index is calculated based on answers by the caregiver to the following questions: 1. Do you answer your child's questions? 2. Do you think boys can play with dolls? 3. Do you think girls can play with cars? 4. Do you think children start talking by nature? 5. Does the child's brain develop from gestation or when the child starts going to school? Each outcome variable is standardized using the mean and standard deviation for control households. The rows shows ITT estimates. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and dummy variable for a set of households surveyed in 2016, as well as village fixed effect. Standard errors reported in parentheses. p<0.1, p<0.05, p<0.01

TABLE A13: Impact of leader's exposure on individual ECD tests

	Leader ITT	**S.e**	**P-Value**	**Obs**
	(1)	(2)	(3)	(4)
Younger children (12 - 36 months)				
ECD index (first principal component)	-0.02	(0.05)	0.68	624
Denver Personal-social	-0.03	(0.07)	0.70	639
Denver Language	0.03	(0.06)	0.56	633
Denver Fine Motor	-0.05	(0.08)	0.49	630
Denver Gross Motor	-0.06	(0.07)	0.42	629
Older children (37 - 84 months)				
ECD index (first principal component)	-0.14	(0.05)	0.01	1,427
Denver Personal-social	-0.02	(0.05)	0.72	1,449
Denver Language	-0.14	(0.05)	0.01	1,449
Denver Fine Motor	-0.01	(0.05)	0.85	1,449
Denver Gross Motor	-0.09	(0.06)	0.15	1,449
Memory	-0.07	(0.05)	0.19	1,449
TVIP (Vocabulary)	-0.13	(0.05)	0.02	1,449
Self control	-0.08	(0.06)	0.17	1,427

Every row corresponds to a separate estimation (equation 1). All outcome variables are standardized using the mean and standard deviation of control households. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variables for a set of households surveyed in 2016, as well as community fixed effects. Standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE A14: Impact of leaders' exposure on ECD outcomes, by leaders' education levels

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT Leader	-0,14**	-0,01	-0,18***
	(0,05)	(0,09)	(0,06)
ITT Leader X Leaders medium educ	0,12	-0,03	0,18*
	(0,08)	(0,14)	(0,10)
ITT Leader X Leaders high educ	-0,03	-0,00	-0,06
	(0,09)	(0,13)	(0,11)
Observations	2051	624	1427
Adjusted R^2	0,457	0,155	0,539
ITT Leader + ITT Leader X Leaders medium educ	0,79	0,69	0,98
ITT Leader + ITT Leader X Leaders high educ	0,02	0,89	0,01

Note: All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE A15: Impact of leaders' exposure on ECD outcomes, by baseline social interactions about ECD

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT Leader	-0,10**	-0,08	-0,11**
	(0,04)	(0,08)	(0,06)
Social (Had interactions about ECD at baseline)	0,02	-0,08	0,05
	(0,07)	(0,09)	(0,09)
ITT Leader X Social	-0,02	0,14	-0,06
	(0,08)	(0,12)	(0,10)
Observations	2051	624	1427
ITT Leader + ITT Leader X Social	0,07	0,48	0,04

Note: Social is a binary variable indicating that the household had talked to at least one other community member about ECD in the week before the baseline survey. At baseline, households were asked whether they had talked to other community members about ECD practices in the last 7 days. This includes interaction about ECD practices with the health promoter, pre-school teacher, primary school teacher, other teachers, family members, neighbours, or elected leaders in the village. About half of the households had talked to at least one other community member about ECD practices. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old, and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE A16: Impact of leaders' exposure on intermediary outcomes, by baseline social interactions about ECD

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
ITT Leader	-0.14	-0.04	-0.06	-0.19**	-0.23***	-0.08	-0.14*
	(0.10)	(0.06)	(0.06)	(0.08)	(0.09)	(0.07)	(0.08)
Social (Had interactions about ECD at baseline)	0.05	-0.06	-0.01	0.03	-0.04	0.06	0.02
	(0.08)	(0.09)	(0.08)	(0.07)	(0.10)	(0.09)	(0.08)
ITT Leader X Social	0.01	0.13	0.13	0.06	0.01	0.18*	0.06
	(0.10)	(0.11)	(0.08)	(0.09)	(0.11)	(0.10)	(0.09)
Observations	2062	2062	2073	2059	2062	1970	2068
ITT leader + ITT leader X Social	0.13	0.35	0.21	0.10	0.07	0.29	0.34

Note: Social is a binary variable indicating that the household had talked to at least one other community member about ECD in the week before the baseline survey. At baseline, households were asked whether they had talked to other community members about ECD practices in the last 7 days. This includes interaction about ECD practices with the health promoter, pre-school teacher, primary school teacher, other teachers, family members, neighbours, or elected leaders in the village. About half of the households had talked to at least one other community member about ECD practices. All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE A17: Impact of leaders' exposure on individual intermediary outcomes: nutrition and stimulation

	ITT Leaders	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Nutrition					
Nutrition index	-0.14	(0.08)	0.08	0.16	2,062
Beverages and soup					
Soy Milk	-0.05	(0.09)	0.60	0.44	2,062
Juice	-0.31	(0.19)	0.11	4.05	2,062
Soup	-0.06	(0.08)	0.44	0.76	2,062
Vegetables					
Salad	0.06	(0.04)	0.17	0.33	2,062
Potatos	-0.19	(0.11)	0.07	1.29	2,062
Tomate	0.02	(0.23)	0.95	3.17	2,062
Onion	-0.10	(0.18)	0.57	2.28	2,062
Other Vegetables	-0.07	(0.13)	0.59	1.75	2,062
Staples					
Rice	0.06	(0.17)	0.74	5.90	2,061
Beans	-0.10	(0.10)	0.29	6.16	2,062
Bread	-0.05	(0.25)	0.83	4.26	2,062
Tortilla	0.11	(0.15)	0.45	5.82	2,061
Cookies	-0.24	(0.12)	0.05	1.53	2,062
Proteins					
Protein index	-0.23	(0.09)	0.14	0.00	2,062
Milk	-0.32	(0.24)	0.19	1.97	2,062
Eggs	-0.23	(0.14)	0.11	2.03	2,062
Cheese	-0.55	(0.24)	0.02	2.48	2,061
Meat	-0.22	(0.12)	0.08	1.37	2,062
Breast milk	-0.18	(0.34)	0.60	3.33	647
Stimulation					
Stimulation index	0.02	(0.06)	0.80	0.13	2,062
Buy a toys for child	-0.01	(0.02)	0.71	0.89	2,062
Tell tales to child	0.02	(0.02)	0.40	0.82	2,062
Read books to child	-0.01	(0.03)	0.84	0.31	2,062
Have paper and pen	0.01	(0.02)	0.73	0.84	2,062

Every line corresponds to a separate estimation (equation 3) and shows the ITT estimates on aggregate indices or individual questions measuring investments in nutrition or stimulation as reported by the caregiver. Questions on food items measure how many days in the last 7 days the child was given the specific item. The nutrition index was calculated by summing the days over all food items. The protein index was calculated by summing the days over all protein items. The individual items about stimulation are answers to yes/no questions, and the index is obtained by summing over yes answers. The nutrition, protein, and stimulation indices are standardized using the mean and standard deviation for control households. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

TABLE A18: Impact of leaders' exposure on individual intermediary outcomes: preventive health

	ITT Leaders	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Health					
Health index	-0.00	(0.04)	0.37	0.00	2,073
Insist on brushing teeth	-0.00	(0.01)	0.78	0.02	2,073
Ensure a healthy diet	0.01	(0.01)	0.05	0.06	2,073
Keep child away from the stove	-0.00	(0.00)	0.88	0.00	2,073
Ensure good hygiene	-0.01	(0.02)	0.27	0.21	2,073
Wash clothes properly	-0.00	(0.01)	0.71	0.12	2,073
Avoid giving coffee	0.00	(0.00)	0.29	0.00	2,073
Make child wash hands	0.00	(0.02)	0.90	0.27	2,073
Smoke at a distance from child	-0.00	(0.00)	0.74	0.00	2,073
Do not give child unhealthy snacks	0.00	(0.01)	0.72	0.01	2,073
Cover fresh food	0.02	(0.02)	0.86	0.13	2,073
Teach child to chew slowly	0.00	(0.00)	0.80	0.00	2,073
Prevent child from catching the flu	-0.01	(0.01)	0.64	0.06	2,073
Boil or chlorinate the water	0.00	(0.01)	0.42	0.02	2,073
Ensure that child vaccines are up-to-date	-0.00	(0.01)	0.34	0.07	2,073
Use a mosquito net	-0.00	(0.01)	0.39	0.01	2,073
Insist child wears shoes	-0.01	(0.01)	0.60	0.03	2,073
Make sure child eats regularly	-0.00	(0.00)	0.46	0.00	2,073
Keep child away from dangerous products	-0.00	(0.00)	0.85	0.00	2,073
Show child affection	0.00	(0.00)	0.26	0.01	2,073
Other	0.01	(0.02)	0.87	0.17	2,073

Each row corresponds to a separate estimation (equation 3) and shows the ITT estimate on the aggregate health index or on individual questions measuring whether the caregiver reported a given practice when asked how to avoid their young child getting sick. The health index is calculated based on the number of different preventive health practices. The index is standardized using the mean and standard deviation of the control household. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. Five villages without leaders were excluded from the estimation. The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE A19: Impact of leaders' exposure on individual intermediary outcomes: micronutrients, ECD attitudes and observed hygiene

	ITT Leaders	S.e	P-Value	Mean Control	Obs
	(1)	(2)	(3)	(4)	(5)
Micronutrients					
Micronutrients index	-0.16	(0.06)	0.02	0.13	2,059
Received Vitamin A	-0.06	(0.03)	0.05	0.76	2,059
Received deworming medecine	-0.05	(0.03)	0.05	0.80	2,059
Received iron supplementation	-0.04	(0.03)	0.15	0.55	2,059
ECD Attitudes of main caregiver					
Attitudes index	-0.00	(0.06)	0.97	0.08	1,970
Should always answer child's questions	-0.01	(0.02)	0.63	0.89	1,970
Boys can play with dolls	0.02	(0.03)	0.62	0.39	1,970
Girls can play with cars (toy)	0.01	(0.03)	0.82	0.53	1,970
Early language develops from stimulation	0.01	(0.02)	0.67	0.68	1,970
Brain develops from very early on	-0.03	(0.01)	0.04	0.88	1,970
Hygiene (observed by test administrator)					
Hygiene index	-0.11	(0.06)	0.09	0.06	2,068
Child has clean face	-0.03	(0.03)	0.25	0.70	2,068
Child has clean hair	-0.05	(0.02)	0.02	0.82	2,068
Child has clean hands	-0.07	(0.03)	0.01	0.61	2,068
Child has clean clothes	-0.06	(0.03)	0.03	0.61	2,068
Child is wearing shoes	0.03	(0.02)	0.20	0.30	2,068
Child does not cough	0.01	(0.02)	0.64	0.82	2,068
Child has clean nose	0.00	(0.02)	0.94	0.83	2,068
Child does not have skin problems	0.00	(0.02)	0.81	0.92	2,067

Every row corresponds to a separate estimation (equation 3). The vitamin index is based on three variables: the child has received vitamin A, ferrous sulfate and deworming medicine during the last six months (yes/no questions). The hygiene index captures the hygienic condition of the child observed by the test administrator. The Attitudes index is calculated based on answers by the caregiver to the following questions: 1. Do you answer your child's questions? 2. Do you think boys can play with dolls? 3. Do you think girls can play with cars? 4. Do you think children start talking by nature? 5. Does the child's brain develop from gestation or when the child starts going to school? Each outcome variable is standardized using the mean and standard deviation for control households. The rows shows ITT estimates. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age, gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. Five villages without leaders were excluded from the estimation. The standard errors (in parentheses) are clustered at village level. $p<0.1$, $p<0.05$, $p<0.01$

TABLE A20: Impact of text messages on ECD outcomes for children of leaders versus non-leaders

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT	-0.01	-0.02	-0.01
	(0.03)	(0.07)	(0.04)
Child of Leader	0.10	0.24*	0.03
	(0.07)	(0.12)	(0.08)
ITT X Child of Leader	-0.06	-0.02	-0.08
	(0.08)	(0.13)	(0.10)
Observations	2393	747	1646
ITT + ITT X Child of Leader =0	0.39	0.73	0.32

Note: All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months old, and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 to 35 months. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, Memory and self-control) of children aged 36 to 83 months. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Five villages without leaders were excluded from the sample. Standard errors reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE A21: Impact of text messages on intermediary outcomes for children of leaders versus non-leaders

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
ITT	0.08	0.19***	0.04	0.07	0.06	0.12**	0.09*
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
Child of Leader	-0.14	0.31***	-0.04	0.21*	-0.03	0.21*	0.05
	(0.11)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.11)
ITT X Child of Leader	-0.09	-0.16	0.15	-0.11	-0.24**	0.06	-0.07
	(0.12)	(0.12)	(0.11)	(0.12)	(0.11)	(0.12)	(0.12)
Observations	2407	2408	2419	2403	2408	2303	2413
ITT + ITT X Child of Leader = 0	0.88	0.82	0.06	0.71	0.09	0.12	0.91

Note: All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. All the regressions include controls for the stratification variables, child age and gender, enumerator fixed effects, and a dummy variable for a set of households surveyed in 2016, as well as village fixed effects. Five villages without leaders were excluded from the sample. Standard errors reported in parentheses.* p<0.1, ** p<0.05, *** p<0.01

TABLE A22: Impact of leaders' exposure on ECD outcomes, by closest leader's predicted commitment

	(1) ECD full sample	(2) ECD young	(3) ECD Old
ITT Leader	-0.05	-0.04	-0.05
	(0.09)	(0.18)	(0.09)
ITT leader X low commitment	-0.13	-0.06	-0.19
	(0.11)	(0.18)	(0.12)
ITT leader X medium commitment	-0.15	-0.12	-0.19
	(0.12)	(0.24)	(0.12)
ITT leader X high commitment	0.08	0.30	-0.01
	(0.11)	(0.20)	(0.13)
ITT leader X very high commitment	-0.05	-0.02	-0.06
	(0.12)	(0.28)	(0.12)
Observations	2051	624	1427
ITT leader + ITT leader X low commitment = 0	0,01	0.33	0,01
ITT leader + ITT leader X medium commitment = 0	0.01	0.20	0.01
ITT leader + ITT leader X high commitment = 0	0.74	0,02	0.54
ITT leader + ITT leader X very high commitment = 0	0.24	0.74	0.19

Note: Low, medium, high and very high commitment indicates whether leader is predicted to be in the 2, 3, 4, or top quintile of leaders with highest participation in the quizzes. All outcome variables are standardized using the mean and standard deviation of control households. In column 1, the dependent variable is the first principal component of the 4 Denver sub-components for children below 36 months and the first principal component of all 7 tests for children 36-83 months old. In column 2, the dependent variable is the principal component for the four Denver sub-components for children aged 12 and 35 month. In column 3, the dependent variable is the first principal component of all 7 tests (4 Denver sub-components plus TVIP, memory and self-control) of children aged 36 to 83 months. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level. * p<0.1, ** p<0.05, *** p<0.01

TABLE A23: Impact of leaders' exposure on intermediary outcomes, by closest leader's predicted commitment

	(1) Nutrition	(2) Stimulation	(3) Health	(4) Micronutrients	(5) Proteins	(6) Attitudes	(7) Hygiene
ITT Leader	-0.08	0.06	0.03	-0.07	-0.01	0.04	-0.04
	(0.15)	(0.09)	(0.10)	(0.12)	(0.14)	(0.10)	(0.11)
ITT leader X low commitment	-0.13	-0.07	-0.06	-0.20	-0.31*	-0.02	-0.21
	(0.16)	(0.15)	(0.15)	(0.16)	(0.18)	(0.15)	(0.15)
ITT leader X medium commitment	-0.18	-0.16	-0.21	-0.17	-0.36*	-0.14	-0.14
	(0.20)	(0.15)	(0.13)	(0.16)	(0.21)	(0.15)	(0.15)
ITT leader X high commitment	0.03	0.01	0.08	0.12	-0.27	-0.21	0.04
	(0.20)	(0.14)	(0.14)	(0.17)	(0.19)	(0.16)	(0.15)
ITT leader X very high commitment	-0.01	0.00	0.03	-0.18	-0.16	0.15	-0.03
	(0.22)	(0.15)	(0.13)	(0.16)	(0.19)	(0.16)	(0.18)
Observations	2062	2062	2073	2059	2062	1970	2068
ITT leader + ITT leader X low committed = 0	0.01	0.89	0.72	0.01	0.045	0.84	0.04
ITT leader + ITT leader X medium committed = 0	0.08	0.34	0.05	0.02	0.03	0.41	0.13
ITT leader + ITT leader X high committed = 0	0.64	0.55	0.20	0.71	0.04	0.23	0.95
ITT leader + ITT leader X very high committed = 0	0.63	0.64	0.42	0.02	0.19	0.13	0.65

Note: Low, medium, high and very high commitment indicates whether leader is predicted to be in the 2, 3, 4, or top quintile of leaders with highest participation in the quizzes. All outcome variables are standardized using the mean and standard deviation of the control household. The first 6 columns show ITT estimates on indices aggregating information regarding early childhood investments, as reported by the caregiver. Column 7 aggregates information regarding hygiene of the child, as observed by the enumerator. All regressions include controls for the household-level treatment, the average level of education of the leaders (the stratification variable for village level leader randomization), as well as the level of education of the main caregiver, whether there is a male caregiver in the household, whether the leader household had access to electricity, child age and gender and test-administrator fixed effects as well as a binary indicator indicating the data was collected in 2016. The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

TABLE A24: Impacts on social interactions about ECD with other members of the community

	(1) # other members	(2) Relatives	(3) Neighbors	(4) Religious leader
ITT	0.05	0.03	0.03	-0.01
	(0.04)	(0.02)	(0.02)	(0.02)
ITT Leader	0.04	-0.01	0.02	0.03*
	(0.05)	(0.02)	(0.03)	(0.02)
Observations	2073	2073	2073	2073
Mean Control	0.95	0.40	0.36	0.19

Note: The estimation corresponds to the estimator β_2 in equation 3. The dependent variables are the social interactions between the head of the household and social leaders during the previous week. Column 1 is an index accounting for household head interaction with relatives, neighbors and religious leader. The remaining columns are binary: 1 whether the household head had at least one interactions with any respective leader during the referent week. All the regressions include controls for the stratification variables: education of the household, titular of the household is male, whether the leader has access to electricity, targeted child age and sex, enumerator and a dummy variables for a set of households surveyed in 2016 (one year after the intervention). The standard errors (in parentheses) are clustered at village level.* p<0.1, ** p<0.05, *** p<0.01

A1.2 Lasso prediction of quiz participation

Information on quiz participation can be seen as a measure of compliance or engagement with the text message intervention. As the control group did not received text messages, they did not participate in the quiz. We therefore use baseline observables to predict the frequency of participation in the quiz for each leader household, using Lasso (Tibshirani (1996)).

We minimize the sum of the square residuals (SSR) correcting with a Lasso regression penalty given by the following expression

$$SSR + \lambda * (\mid \beta_1 + \beta_2 + \beta_3 + ... + \beta_k \mid)$$

where lambda is the penalized Lasso parameter, and β_k are the OLS parameters for the variables k, excluding the intercept.

When $\lambda = 0$ the Lasso prediction will be the same as the OLS prediction. But as λ increases, Lasso will shrink the slope of a subset of parameters to 0, reducing the number of covariates used for the prediction.

The Lasso regression has more bias than OLS, but it has lower variance of the predicted values and hence improve the overall prediction accuracy.

To obtain the predictive model, we start from 145 baseline variables, including questions on parental practices, nutrition and stimulation of the child, household economic activity, migrant status, social interactions with the rest of the community, labor supply of all household members, as well as a demographic characteristics of the household head and targeted child. We also included questions on education and health. We force the model to keep all stratification variables.

By setting lambda = 10, we can explain 65% of the variation keeping the following 22 variables[31]: caregiver's assessment of child's' fine motor skills, household is engaged in agriculture activity, has livestock activities, and has private wage job activities, the head of the household has a small manufacturing business activity, someone in household has migrated to work as a nanny, someone in the household has received training on ECD,

[31] By comparison the OLS model (lambda =0), leads to a model with $r^2 = 0,662$.

the household has had social interactions with the health promoter, with the primary school teacher and a local political leader, the caregiver considers that hitting a child is a good parenting practice, the caregiver considers that playing games and giving affection is good parenting practice, someone in the household sometimes threatens the child by saying that a monkey may take him away, someone in the household sometimes uses a phone in other households, someone in household has a cell phone with Movistar coverage, the household has Movistar network coverage at home, the child drank milk during the previous week, number of days that the child consumed fruit in previous week, at least one child has a caregiver other than his/her mother, household size, number of adults in the household who have migrated temporarily over last 12 months, household has access to electricity, and presence of a male caregiver.