Journal Pre-proof

Can mHealth campaigns improve CCT outcomes? Experimental evidence from sms-nudges in Indonesia

Friederike Lenel, Jan Priebe, Elan Satriawan, Ekki Syamsulhakim

PII: S0167-6296(22)00102-3
DOI: https://doi.org/10.1016/j.jhealeco.2022.102687
Reference: JHE 102687

To appear in: *Journal of Health Economics*

Received date : 14 January 2021
Revised date : 5 September 2022
Accepted date : 13 September 2022

Please cite this article as: F. Lenel, J. Priebe, E. Satriawan et al., Can mHealth campaigns improve CCT outcomes? Experimental evidence from sms-nudges in Indonesia. *Journal of Health Economics* (2022), doi: https://doi.org/10.1016/j.jhealeco.2022.102687.

Can mHealth campaigns improve CCT outcomes? Experimental evidence from sms-nudges in Indonesia*

Friederike Lenel* Jan Priebe† Elan Satriawan# Ekki Syamsulhakim‡

September 5, 2022

Employing a clustered RCT this paper examines the short-term impact of a 12-month mHealth BCC campaign on health practices & outcomes among CCT beneficiaries in Indonesia. Our analysis reveals that the intervention led to substantial improvements in maternal health behavior (postnatal care, child vaccinations, hygiene practices) & outcomes (anemia rates). Adopting a heterogeneous treatment effect framework, we further show that improvements in maternal hygiene practices and anemia rates are closely linked to health knowledge gained by mothers through the sms campaign. In contrast, we provide suggestive evidence that improvements in other health indicators are more likely to be related to the reminder and nudge components of the intervention.

Key words: CCT, mHealth, information, Indonesia.

***JEL* codes:** D80, I10, I38, J13.

*We would like to thank the Development Economics Research Groups at BNITM, DENeB, and the University of Göttingen for valuable comments. The paper benefited from discussions with Jenny Aker, Macartan Humphreys, Stephan Klasen, Benjamin A. Olken, and Sudarno Sumarto. We are grateful to Vivi Yulaswati from Bappenas, Suahasil Nazara from Ministry of Finance, Bambang Widianto from TNP2K, and Changqing Sun from the World Bank for their support. Excellent work in the implementation of the intervention was provided by Airin Roshita, Kartika Sari Juniwaty, Rizal Adi Prima, and Rizky Ika Syafitri. We thank the Australian government's Department of Foreign Affairs and Trade (DFAT) for financial support. The views expressed in this paper are those of the authors alone and do not present the views of BNITM, TNP2K, and the World Bank. Declarations of interest: none. IRB approval was obtained from the University of Indonesia.

* Department of Economics, University of Göttingen.

† Corresponding author: Bernhard Nocht Institute for Tropical Medicine, Bernhard Nocht Straße 74, 20359 Hamburg, Germany. Email: jan.priebe@bnitm.de.

TNP2K & Universitas Gadjah Mada.

‡ World Bank & Universitas Padjadjaran.

1. Introduction

Cash transfer programs have been associated with a wide range of positive outcomes, including poverty reduction, higher business incomes, and increased schooling (Bastagli et al., 2016; Fiszbein et al., 2009). The evidence with respect to maternal and child health, however, is less conclusive with recent reviews and meta-studies pointing out that some evaluations have shown positive impacts, while several others reported mixed or null results (Baird et al., 2014; Bastagli et al., 2016; Biscaye et al., 2017; Gaarder et al., 2010; Ladhani and Sitter, 2020; Millan et al., 2019; Tirivayi et al., 2016).

To strengthen the impact of cash transfers on health outcomes, these programs have begun to be re-designed as cash-plus programs, combining cash with additional interventions, services, or transfers (Bhutta et al., 2013; Glassman et al., 2013; Roelen et al., 2017). In terms of improving maternal and child health outcomes, a popular complementary component involves behavioral change communication (BCC), an approach that emphasizes the role of information in health production by promoting certain knowledge, attitudes, beliefs, and practices (Avula et al., 2013; Carneiro et al., 2021). While in general the evidence on the effectiveness of cash-plus BCC components appears still to be scarce and inconclusive (Gilligan et al., 2020; Little et al., 2021; Premand and Barry, 2022), some recent studies show that child nutrition outcomes (Ahmed et al., 2019; Field and Maffioli, 2021; Khan et al., 2019) and maternal knowledge (Levere et al., 2016) could only be improved in treatment arms that involved complementary BCC components.

The 'traditional' implementation of BCC interventions, as studied in the existing health-related cash-plus literature, typically involves regular individual and/or group-based face-to-face meetings with program facilitators. In this context, given high operational costs, 'traditional' BCC interventions often need to be strictly targeted towards individuals who are (i) particularly vulnerable or (ii) more likely to contribute to spreading and disseminating information within their communities. Consequently, adding a -– possibly beneficial – 'traditional' BCC component to cash transfers is not a trivial policy decision as it typically involves new implementation challenges and leads to substantially higher operational costs of an often already expensive program (Benhassine et al., 2015). In particular, in a context in which governments have limited resources for social protection the extra costs involved for recruiting and training qualified staff, travel expenses, and wages of facilitators need to be carefully considered.

A promising alternative to the more 'traditional' BCC implementations relates to low-cost mHealth solutions. Systematic reviews have shown that mHealth interventions are feasible across a wide range of cultural and geographical settings with a number of studies showing positive impacts of standalone mHealth BCC interventions.[1] Furthermore, given the comparatively low cost of mHealth campaigns, these interventions are more likely to be offered universally and do not necessarily require stricter access criteria and targeting as part of its implementation.

Ultimately, it is an empirical question whether an mHealth BCC component can be

[1] See for example Badawy et al. (2017); Celik and Toruner (2020); Crutzen et al. (2011); Gurman et al. (2012); Majeed-Ariss et al. (2015); Loescher et al. (2018); Palmer et al. (2018); Payne et al. (2015).

successful in the cash-plus context. First, it is difficult to extrapolate findings from existing mHealth BCC studies. Besides documented large variations in effect sizes, the circumstance that the target population is enrolled in a cash program modifies several parameters such as attentiveness, reciprocity, and trust relationships, which have been associated with the performance of mHealth interventions. Second, insights and benchmarks gained from the existing cash-plus BCC literature cannot be readily transferred to the mHealth context as intervention features differ in important ways.

In this paper, we examine whether a simple and low-cost mHealth BCC component can improve health behavior & outcomes among poor women who are enrolled in a conditional cash transfer program (CCT). Specifically, our research attempts to answer three main questions: Can a text messaging campaign help increase maternal health practices and outcomes among CCT beneficiaries? Which type of health indicators are particularly responsive? Which implementation and behavioral factors are correlated with the success and/or failure of such a campaign?

The context of our study is PKH (Program Keluarga Harapan) in Indonesia. PKH covers about 10 million households and is the country's flagship social protection program. Evaluations of PKH revealed that the program led to poverty reductions, increases in school enrollment (Cahyadi et al., 2020; WB, 2012), and a mixed impact on a broad range of health indicators (Christian et al., 2019; Kusuma et al., 2016; Triyana, 2016). While PKH managed to reduce child stunting rates by 10pp in the long-term, the program did not show lasting improvements in health knowledge, breastfeeding practices, pre- and post-natal care, and childhood vaccinations (Cahyadi et al., 2020).

In response to PKH's limited impact on maternal and child health, we collaborated with the government of Indonesia (GoI) to explore whether an mHealth BCC campaign—more specifically, one-directional text messages—can improve maternal health knowledge, behavior, and outcomes among PKH beneficiaries in relation to five health topics: postnatal care, child vaccinations, hygiene practices, maternal anemia, and child anthropometrics. The intervention involved sending out three health-related sms per week over a period of twelve months. To evaluate the short-term impact of the intervention, we implemented a clustered-RCT in which 1,821 beneficiaries in 127 villages were randomly assigned to either a control group (PKH) or a treatment group (PKH + mHealth BCC).

Overall, our results suggest that the mHealth BCC campaign substantially improved maternal health behavior & outcomes. Mothers' postnatal care practices improved by 8pp (15%), child vaccination rates by 7pp (9%), hygiene practices by 10pp (50%), while maternal anemia rates dropped by 10pp (20%). Regarding child anthropometric outcomes, we obtain suggestive evidence that the sms campaign might have helped reduce stunting rates among children below the age of five—in particular children who were exposed to the intervention at early ages (2-14 months of age at baseline). In contrast, we show that the sms campaign did not improve the incidence of wasting among children under five.

Examining possible channels that may help explain our main findings, we demonstrate that the sms campaign successfully improved mothers' health knowledge in three out five targeted health dimensions. Adopting a heterogeneous treatment effect framework, we

further provide suggestive evidence that particular implementation design features affect maternal health behavior & outcomes differently. While we find that health knowledge gains are strongly associated with improvements in hygiene practices and anemia reduction, we believe that the reminder and/or nudge components inherent in our sms campaign are more likely to explain observed changes in postnatal care practices and child vaccination rates. Lastly, employing a behavior change theory framework we illustrate that our interventions seems to have been particular successful among mothers who had a positive attitude towards the intervention.

Our experiment advances the relevant literature in two ways. First, we add to the existing cash-plus literature by examining the effectiveness of complementary components that leverage mobile devices and technologies. While the existing literature covers a broad set of different types of interventions related to the role of conditionalities (Baird et al., 2011; Benhassine et al., 2015; Brollo et al., 2017), benefit amounts (Filmer and Schady, 2009; Haushofer and Shapiro, 2016), benefit types (Ahmed et al., 2019; Langendorf et al., 2014), access to services (Banerjee et al., 2021; Guanais, 2013), and in-person face-to-face interventions for trainings, mentoring, and counseling (Ahmed et al., 2019, 2020; Arriagada et al., 2020; Fiala, 2018; Gilligan et al., 2020; Karlan et al., 2015; Premand and Barry, 2022; Roy et al., 2019; Sedlmayr et al., 2020), only little evidence on the impact of mobile solutions yet exists.[2] Our findings suggest that mobile solutions can be an effective complement and/or substitute of more conventional in-person interventions.

Second, we contribute to the mHealth literature in economics and public health. To the best of our knowledge existing studies in economics have almost exclusively focused on two subgroups that are very different from the population we study: (i) patients (Pop-Eleches et al., 2011; Raifman et al., 2014) and (ii) the general population (Bahety et al., 2021; Banerjee et al., 2020; Dammert et al., 2014; Falco and Zaccagni, 2021; Grepin et al., 2019).[3] In addition to these studies, we examine a different and much broader set of health indicators.

Furthermore, while there are plenty of studies in the medical and public health literature that investigate the impact of mHealth interventions, systematic reviews and meta-studies have pointed out that only few studies go beyond intention-to-treat esti-

[2] A notable exception is Grepin et al. (2019) in which the authors examine the role of sms reminders on maternal health (safe deliveries) in Kenya. The authors find that sms reminders can increase the impact of cash transfers. Since the cash transfers in Grepin et al. (2019) were handed out directly by the researchers, involved strict following of conditionality criteria, and did not involve any poverty targeting, our study provides a proof-of-concept for a public sector implementation context.

[3] Pop-Eleches et al. (2011) find that HIV patients in Kenya are more likely to adhere to treatment as a result of a text messaging campaign, while Raifman et al. (2014) observe that malaria patients in Ghana are more likely to adhere to treatment when receiving text message reminders. With regard to non-patient studies, Banerjee et al. (2020) and Falco and Zaccagni (2021) find that the framing of text messages matters in people's compliance with COVID-19 related health and social distancing behaviour in Denmark and India. In contrast, Bahety et al. (2021) find that sms messages did neither increase COVID-19 related knowledge, nor social distancing and handwashing practices in India. Furthermore, Dammert et al. (2014) find that sms notifications reduce transmission risk of dengue by 0.1SD in Peru, while Grepin et al. (2019) find a very limited effect of sms nudges on pregnant women's likelihood to deliver in health facilities in Kenya.

mates with results often somewhat appearing questionable due to inadequate controls for selection effects (into and out of the intervention/survey) and small sample sizes (Agravat, 2013; Blaya et al., 2010; Cole-Lewis and Kershaw, 2010; Fjelsoe et al., 2009; Hall et al., 2015; Orr and King, 2015; Sondaal et al., 2016). Moreover, as described in Armanasco et al. (2017) there is a lack of mHealth studies that is able to elicit pathways and mechanisms through which mHealth interventions work. By employing a research framework that links to mHealth implementation features (information vs. reminders vs. nudges) and behavior change theories (theory of planned behavior), we believe that our study is able to shed more light on contextual factors that facilitate or prevent the success of mHealth interventions.

This paper proceeds in six sections. Section II provides background on PKH and our intervention. Section III describes our data and field implementation. Section IV presents the main results and explores robustness checks. Section V examines moderators and pathways responsible for our results. Section VI discusses policy implications related to the main findings. Section VII offers concluding thoughts.

2. Background

2.1. The Conditional Cash Transfer Program: PKH

Program Keluarga Harapan (PKH) was introduced in 2007 for 600,000 households and subsequently expanded to cover 10 million households in 2019 (MoSA, 2020). At the end of 2014 – the time of our baseline survey – PKH served about 3 million households across Indonesia. To be eligible for PKH, a household has to be poor[4] and in addition to fulfill at least one of the following demographic criteria: at least one child in the household is below the age of 16 and/or at least one woman in the household is pregnant. PKH provides sizeable cash transfers. Households receive between 83 and 290 US dollars per year depending on a household's demographic structure.[5] On average the transfer constitutes about 15 percent of annual household expenditures of a poor household.[6]

PKH comprises five conditionality criteria related to health.[7] Three criteria refer to

[4]The poverty status is determined by proxy-means tests which are linked to the country's national targeting database. During our study period a household was considered eligible for PKH if it belonged to the poorest 8 percent of all Indonesian households. Compared to other social assistance programs in the country, PKH is considered to be well targeted (Alatas et al., 2019).

[5]The average payment was 187 US dollars at 2013 prices (Nazara and Rahayu, 2013; WB, 2017). PKH payment structures in 2014 provided a base payment to every eligible household of about 30 US Dollars (300,000 Indonesian Rupiah), with different top-ups depending on the number and age of children and whether the mother was pregnant.

[6]According to Indonesia's socio-economic survey (SUSENAS), the poorest 10 percent of households spent in 2014 on average about 14 million Rupiah (1,400 US dollars) per year.

[7]In practice, however, the monitoring, verification, and enforcement of the health conditionality criteria has been a constant implementation challenge and was occasionally abandoned. Albeit there exist large regional and facilitator-specific variations in whether and how strict conditionality criteria are enforced, in most cases violations of conditionality criteria are not punished. Furthermore, only a fraction of PKH beneficiaries were found to strictly follow PKH health conditionality criteria (MSC, 2019; WB, 2012).

maternal health and consist of the following: 1.) complete four antenatal care visits and take iron pills during pregnancy, 2.) be assisted by a trained professional during birth, and 3.) mothers with infants must complete two post-natal care visits. Two conditionality criteria apply to young children (< 6 years): 4.) ensure that children have complete childhood immunization and take vitamin A capsules twice a year, and 5.) take children to regular growth monitoring check-ups.[8] All five conditionality criteria are expected to be served by local level health staff and infrastructure.

At the local level PKH implementation is supported by a facilitator who—depending on the location—is responsible for 50-80 PKH households. The principal tasks of facilitators are administrative and involve the provision of PKH-related information (including health topics), monitoring of conditionalities, and organization of so called PKH groups (kelompok PKH). PKH groups consist of 10-25 beneficiaries (mothers) and are led by the so called PKH group mother (ibu kelompok PKH), who is a PKH benefiaciary herself.[9] According to program guidelines, PKH groups convene every month in order to discuss PKH related topics such as benefit payments as well as education and health topics.

2.2. Intervention Set-Up

Rational The GoI continuously aims at strengthening PKH in order to boost health knowledge, practices, and outcomes among PKH beneficiaries. The country's context with almost universal cell phone penetration (with respect to signal and ownership) in combination with minimal sms delivery costs make mHealth and in particular text messaging interventions a very attractive option to be tested.

Furthermore, PKH routinely stores cell phone numbers from each beneficiary. Prior to our intervention, phone numbers of PKH households were used by PKH administrators to communicate the timing of benefit payments. Likewise, PKH faciliators occasionally used the phone numbers to collect information about and coordinate visits of PKH households. Therefore, PKH households were already familiar with receiving PKH-related information via mobile phone including sms. Additionally, the almost universal cellphone possession among the poor would allow for a possible nationwide roll-out of the mHealth campaign.

The objective of our text messaging campaign was to improve in the short-run health knowledge, behavior, and outcomes of PKH households. The evaluation framework of the intervention (as discussed below) was designed to provide rigorous empirical evidence on short-term effects.

The selection of topics for the mHealth campaign was guided by PKH's objective to boost maternal and child health. Eventually, five main topics were chosen. Anemia was selected since Indonesia exhibits comparatively high rates of anemia in combination with high rates of miscarriages and maternal mortality (DHS, 2018). The remaining four topics concerned breastfeeding, child immunization, hygiene (hand washing), and

[8] Check-ups are supposed to be monthly for infants and quarterly for children 1-5 years old.

[9] PKH groups are formed at the village level. If a village has more than 25 beneficiaries multiple PKH groups are formed in the location.

post-natal care behavior with the latter topic aiming at improving child anthropometric outcomes.

Intervention outline The mHealth BCC intervention was implemented between March 2015 and February 2016 with the aim to boost mothers' health knowledge, behavior, and outcomes. Targeted PKH households were sent three sms per week. In total, each household received 156 sms over a period of 52 weeks. As shown in Table 1 below all five selected health topics featured throughout the campaign.[10]

For the implementation of the intervention several additional steps were taken to increase the chance that PKH households paid attention to the sms. First of all, all messages were personalized so that recipients were more likely to engage in reading the texts (Doss et al., 2019). Second, by partnering with Indonesia's major telecommunication companies and GoI, the sender of the sms was always shown as 'PKH information' in order to create trust into the reliability of the information source. Third, the sms delivery time was adjusted to when mothers would be most likely to have time to read the sms.[11]

Table 1 about here.

The development of the framing for the sms messages was done in consideration of a number of psychological aspects and cognitive biases (Thaler, 2018). In this regard, it was taken into account that the target audience was likely to possess a limited attention span and a high level of stressors and cognitive load due to the circumstance of motherhood, raising at least one young child, and the poverty context with the latter being linked to lower cognitive functioning and resources (Mani et al., 2013; Schilbach et al., 2016).

Consequently, the language used in the sms was adjusted to be non-technical, universal, and easy to understand to facilitate the absorption of information (Fricke et al., 2018; York et al., 2018). Furthermore, the intervention leveraged a mix of sms that contained new information (about 54 percent of all sms) and sms that repeated previously sent messages (reminders; about 46 percent of all sms). Moreover, as shown in Table 1 the majority of sms (58 percent) used gain vs. loss framings as such content, derived from prospect theory, has often been linked to inducing stronger health behavior changes (Fetter et al., 2019). Likewise, about 60 percent of sms related health content to the benefit of a child's (infant) health with the wording - grounded in psychology's

[10] Within a given week all three sms related to the same health topic. Moreover, evidence from meta-studies on sms-nudges in health has shown that the optimal intervention period is six to twelve months (Armanasco et al., 2017) with the optimal number of sms per week being three (Cortes et al., 2020).

[11] Based on research from pre-tests, each week one sms was sent out on Tuesday, Thursday, and Saturday at 7pm — a time when mothers are usually at home, less busy with child care obligations, and tended to access their cell phones. For ease of implementation the text messaging campaign was one-directional. This means that while targeted PKH households received the sms, households could not reply to the sms and could not contact any specific hotline or website.

attachment theory (Fonagy et al., 1991; Slade, 2008) and the theory of the mind (Baron-Cohen, 2000) – chosen to make mothers more comfortable in dealing with their child and stimulating mothers' role as caregivers. Lastly, almost all sms (99 percent) consisted of two parts: (i) health information and (ii) actionable advise. Following recommendations from the mHealth literature (Unicef, 2018) providing direct advise is meant to address inattention and decision fatigue by decomposing complex tasks into simpler ones. A typical sms would look like the example below:

Example immunization:

> Ms. Anindyah, don't wait until the child is sick. Take the child to the Puskesmas for immunization. Children who are immunized are healthier and stronger in facing disease attacks.

The phone numbers were provided by PKH households during the baseline survey (see below). In case multiple cell phones existed in a household, all phone numbers were contacted during the intervention. Monitoring data provided by Indonesia's telecommunication companies suggested that about 99.4 percent of all sms were successfully delivered over the course of the intervention.[12]

3. Experimental Design and Data Collection

3.1. Sampling and Randomization

The initial sampling frame was based on administrative data from PKH's monitoring and information system and comprised 2,400 PKH households in 140 villages in proximity to two urban centres on the islands of Sumatra (city of Pekanbaru) and Sulawesi (city of Makassar).[13] Villages eligible for the study needed to possess a minimum of ten PKH recipients. Otherwise, the 140 villages were drawn randomly from a sampling frame of about 340 villages that were in a radius of two hours of travel time from the respective urban center.

The sample of 2,400 PKH households was restricted to households with a relevant demographic composition for the intervention. Therefore, households had to have at least one child below the age of five. PKH households who only comprised older children were not included in the study.[14]

[12] PKH's standard operating procedure is to constantly update each household's telephone number. Therefore, the occasional switching/churning of phone numbers did not affect the implementation of the intervention.

[13] The GoI wanted to have the pilot implemented in areas which were considered to be understudied but that harbor a substantial share of the Indonesian population. Since most research is typically conducted on the island of Java, Javanese places were excluded from the sampling frame. Sumatra and Sulawesi are the most populated islands after Java. Within Sumatra and Sulawesi the locations were randomly selected conditional on already existing PKH operations in the area.

[14] An exceptions concerns women who were pregnant at baseline and who did not have children yet. These women are already eligible for PKH and were included in our study.

In a given village, each pre-selected household was surveyed (total of 2,400 households). Analysis of the baseline data revealed that the administrative data had not been updated in some cases. More specifically, 579 out of the 2,400 households were dropped from the sample since they did not fulfill the inclusion criteria (no child below the age of five). As a result, the sampling frame for the intervention was reduced to 1,821 households in 127 villages.[15]

The randomization was stratified at the province level and clustered at the village level. Out of the 127 villages (1,821 households) 63 villages (897 households) became part of the control group (PKH) and 64 villages (924 households) became part of the treatment group (PKH + mHealth BCC).

Table C.5 in the Online Appendix C.3 presents estimates of the related minimum detectable effect (MDE) for the principal health indicators used in this study. In terms of percentage point changes (pp) the study design is more likely to detect changes in health knowledge compared to changes in practice indicators and health outcomes.

3.2. Data Collection

The baseline survey was conducted in October and November 2014 with all 2,400 households of the initial sampling frame. In May 2016, about three months after the intervention finished, the endline data was gathered. The endline data collection focused exclusively only on those households that were part of the clustered-RCT (1,821 PKH households).[16] Out of the 1,821 relevant baseline households, 95 could not be interviewed at endline (5%), leaving us with an endline sample of 1,726 households.

The survey targeted as main respondent women and in particular the mother of the children. Besides a standard household roster, the questionnaire comprised modules on the socio-economic background of the household, a comprehensive health module, as well as sections capturing the cognitive ability, personality traits, phone usage behavior, and the household decision-making process of the respondent.

In addition to the interviews, each respondent was asked to provide a blood-sample in order to measure anemia. It total, blood samples were collected from 1,720 PKH mothers at baseline and endline.

Furthermore, the weight and height of the youngest child of the respondent were measured. The collection of anthropometric measures focused on children below the age of 60 months. Since in 228 cases the youngest child was older than 60 months at the time of the endline survey, the final sample comprises 1,492 children.

[15] Please see Figures C.1 to C.5 in the Online Appendix C.1 for maps concerning sample locations.

[16] The data was collected by PUSKA-UI - the mother and child health department at the University of Indonesia.

3.3. Sample Characteristics

3.3.1. Respondent and Village Characteristics

Table A.1 in Appendix A.1 depicts baseline characteristics of villages and respondents included as part of the clustered RCT.[17] Village characteristics, with the exception of PKH related information, are based on Indonesia's village census (PODES) conducted in 2014.

About 40% of villages are located in rural areas with the remaining ones being located in semi-urban areas. On average a village comprises about 3,000 households and for 60% of all villages agriculture constitutes the main source of employment. All villages have cell phone signal, most villages (85%) have a strong signal. All villages have a local health post (posyandu), which focuses on maternal and child health, providing family planning guidance, vaccinations and nutritional supplements. In addition, on average villages have four health centers and two early childhood facilities. Midwives are present in about 67% of villages. On average, villages possess about three PKH groups; one PKH group has on average 14 members.

Almost all respondents are female (99%). Consequently, we refer to respondents as 'mothers' throughout the paper.[18] The majority of respondents are between 15 and 32 years old.[19] Respondents possess on average about 6 years of education, nearly all are married. At baseline, respondents had on average three children. About 9% of sampled women were pregnant at baseline.

3.3.2. Outcome Variables

The analysis in this paper focuses on six health indicators. Three indicators are related to changes in health behavior, namely postnatal care, child vaccinations, and hygiene practices. Each of the three indicators was verified by the enumerator. Postnatal care practice refers to whether mothers kept and used maternal and child health books and could present these books to the enumerator.[20] Vaccination practice indicates whether

[17] Please see Table C.1 and C.2 in the Online Appendix C.2 for the description and coding of village and household variables.

[18] Of all respondents, 3.6% (65 observations) are not the biological mother of the youngest child but another relative who is the main care-taker. This is either because the mother works outside the village and comes home only irregularly, or because the mother has passed away. For 15 observations the respondent is the father; for the remaining cases, the respondent is another female relative, typically the child's grandmother. All our results hold when excluding respondents that are not the biological mothers.

[19] As the sample also includes non-biological mothers, a small proportion of respondents are comparatively old. To account for potential non-linear effects we will use age brackets in the following analyses (age 25 or younger, age 26-35, age above 35).

[20] Postnatal care behavior is affected by the health status of mothers and children. Since we cannot distinguish from the data whether actual postnatal care visits were done because of sickness, precaution, or compliance with PKH, our indicator proxies mothers' diligence towards postnatal care. Our postnatal care indicator (book keeping) is based on verified information. Enumerators asked mothers to show the related books and only if the enumerator was able to verify that the book existed in the household, it was coded as verified. Analyzing the baseline data we find that our measure of

a child was vaccinated at least once as recorded in the child vaccination booklet, which was presented to the enumerator, while hygiene practice is based on stated and observed hand washing practices. In addition, we construct an aggregated health practice index that represents the average over the three behavior related indicators.

Another three indicators are related to health outcomes, namely mothers' anemia and children's stunting and wasting status. Anemia classifications are based on the hemoglobin level measured in the mothers' blood samples taken by enumerators. A mother is defined as having anemia if her hemoglobin level indicates a moderate or severe form of anemia according to WHO standards.[21] Furthermore, the weight and height of the youngest child was measured by enumerators, and children were categorized as stunted or wasted applying WHO child growth reference standards.

Besides the six principal health indicator variables listed above, we additionally investigate in the mechanism section the role of health knowledge in helping to explain possible impacts. The construction of health knowledge indicators follows closely the selected topics for the sms campaign (postnatal care, vaccinations, hygiene, anemia, and breastfeeding). The related variables are indices that describe the share of correct responses by the respondent in a given domain. To gauge knowledge mothers had to respond numerous open-ended questions per topic such as: What are the consequences of anemia? What type of basic immunization should be given to babies? When are you supposed to wash your hands? In addition to the five topic-specific knowledge indices, we construct an aggregate knowledge index which measures the proportion of correct responses across all topics.

The surveys were designed to limit potential biases in the variables of interest. To eliminate or at least reduce the risk of social desirability bias and Hawthorne effects influencing data quality certain precautionary measures were taken. First, all indicators of health behavior are verified by the enumerator. For instance, enumerators asked mothers to show them the child health and vaccination books and inspected toilets/washing areas and handwashing procedures. Only if the respective books could be shown and only if a specific activity/vaccination was listed and signed/stamped by a health professional the information was considered to be verified. Second, indicators for health outcomes are based on measures taken by the enumerators (i.e. the mother's blood sample, the youngest child's weight and height). Third, survey questions on health knowledge differed from the text used in the sms campaign. Consequently, respondents could not simply repeat the sms text when answering knowledge questions, but needed to have processed and understood the information gained during the sms campaign to answer questions correctly. Fourth, given the light-touch nature of the intervention and the circumstance that PKH's standard implementation regularly sends sms to respondents, respondents should not suspect a link between the surveys and the sms campaign.

postnatal care behavior is strongly and positively correlated with mothers' self-reported postnatal health care visits. Results are available from the authors upon request.

[21] According to the WHO standards, a woman has mild anemia with hemoglobin levels between 9.0 and 10.9 g/dL, moderate anemia with hemoglobin levels between 7.0 and 8.9 g/dL, and severe anemia with hemoglobin levels less than 7.0 g/dL.

Table 2 about here.

Summary statistics are shown in Table 2. Panel A depicts variables at baseline for all respondents that were included in the RCT, while Panel B restricts the sample to those that were re-interviewed at endline.[22] With respect to health behavior we observe great variations across indicators. At baseline two out of three mothers do not have any child health record book at home and only 18% own the required two. 75% of the mothers report to have vaccinated their child at least once. Regarding health outcomes we find that at baseline about one out of three mothers has a moderate or severe level of anemia, 38% of the children are stunted and 12% are wasted.

Furthermore, health knowledge varies greatly across different domains and respondents. Mothers seem to have a rather good knowledge of breastfeeding but know little about anemia.

3.3.3. Balance and Attrition

To assess the comparability of observations between the control and treatment group in our analytical sample (1,726 women), we compare the single and joint distribution of variables that could possibly be correlated with our main outcomes.[23] Differences across the control and treatment group might occur due to imbalances stemming from the initial randomization and/or selective attrition. Besides health knowledge and behavior at baseline, we include (i) baseline characteristics of the mother that might affect pre- and post-natal health knowledge and practices as well as health outcomes; (ii) characteristics that might reflect the mother's willingness and ability to adjust beliefs and change behavior over time; and (iii) village characteristics related to health care provision.[24]

Overall, we find that characteristics are balanced between treatment and control, both in terms of respondent and village characteristics (Table A.2 and Table A.3 in Appendix A.1) as well as our outcome variables (Table A.4). However, when testing for joint orthogonality, we have to reject the hypothesis that these characteristics cannot jointly explain treatment status (χ^2=43.17, p-value=0.005); more specifically, conditional on other characteristics, we observe imbalances in baseline postnatal care practice and PKH meeting attendance (Table B.2 in Online Appendix B.1).

We test whether these imbalances are due to non-random attrition. Of the 1,821 respondents, 95 could not be followed-up at endline. While we find that on average there is no differential attrition between treatment and control (see column (1) in Table B.3 in Online Appendix B.1) some characteristics of attriters vary by treatment assignment (see column (3) in Table B.3). In particular, treated mothers who have higher baseline breastfeeding knowledge and a lower number of children in the household are more likely to attrit.

[22] Please note that children's vaccination status was self-reported at baseline, while it was verified by the enumerator at the endline survey. At endline the correlation between the self-reported and the verified vaccination measure was 0.39 (p-value: 0.0004).

[23] The following analysis is guided by Bruhn and McKenzie (2009) and Athey and Imbens (2017).

[24] Table B.1 in Online Appendix B.1 confirms the relevance of these characteristics for health behavior, health knowledge and health outcomes at baseline.

Empirically, we will address the issue of imbalances and selective attrition by including the relevant respondent and village characteristics as controls. Furthermore, as part of our robustness analyses, we will bound our estimated effects varying the assumptions for the outcome values of attriters.

3.4. SMS Handling

To better understand how the sms campaign worked in the field, we analyze self-reported data on sms handling from the endline survey. More specifically, all mothers in the treatment and control group were asked whether in the last year they had received any information about health and nutrition via sms, on which specific topic and from which source. We believe there are three important take-aways.

First, a large share but perhaps not all women in the treatment group received and read our text messages. As shown in Table A.5 Panel A in Appendix A.2, the majority of women in the treatment group (82%) report to have received at least one of the topics covered in the campaign and almost all identify PKH as sender of the message. Moreover, 53% report to have received messages on all topics.

Second, those mothers in the treatment group, who state to have received text messages, report to have found the sms content interesting and helpful. Moreover, they show active engagement with the received texts. While a small proportion of PKH mothers (4%) indicate to have forwarded some of the sms, a majority (64%) report to have discussed the message with others, and 30% discussed the messages in PKH meetings (see Table A.5 Panel B).

Third, we find very little empirical evidence for sizeable spillovers between treatment and control villages. Among PKH mothers in the control group, 6% report to have received messages on at least one of the topics included in our sms campaign, though only 3% identify PKH as source.[25]

4. Main Results

4.1. Empirical Specification

Our empirical identification strategy to estimate treatment effects of the sms campaign rests on two distinct approaches. Under approach 1 we interpret results as intention-to-treat effects (ITT), while under approach 2 we derive local average treatment effects (LATE).

ITT estimates focus on the original treatment assignment. More specifically, we estimate equation 1 below by OLS in which the outcome variable of interest Y of mother i living in village v at time of the endline survey $(t+1)$ is regressed on a treatment

[25] Overall, only four mothers in the control group report to have received messages on all topics. Given that mothers could easily make erroneous claims (e.g. through memorizing sms wrongly) about having received PKH messages, we believe that the results underscore that spillovers are unlikely to play a relevant role in our context.

dummy, T_{vt} that indicates whether a mother has been assigned to the sms intervention based on her village of residence. Equation 1 is specified as follows:

$$\text{ITT} \qquad Y_{ivt+1} = \alpha + \beta T_{vt} + \gamma Y_{ivt} + \delta' X_{ivt} + \zeta' Z_{vt} + \phi_d + \epsilon_{iv} \tag{1}$$

In the following, we estimate two separate specifications of equation 1; one without any covariates and one with covariates. For the model with control variables, we include (i) the lagged dependent variable Y_{ivt}, (ii) individual-level characteristics X_{ivt}, and (iii) village level characteristics Z_{vt}; the covariate selection is guided by the balance and attrition analysis (Section 3.3.3).[26] In addition, we include (iv) sub-district fixed effects ϕ_d.

Approach 2 takes into consideration that compliance with treatment assignment might not have been perfect. For instance, some women in the treatment group claimed that they had not received the message, while a few women in the control group reported to have received the sms. To address this issue we adopt a LATE framework in which we estimate equations 2 by 2SLS. In this set-up we instrument Exp_i (exposure to treatment) by the original treatment assignment (T_{vt}). We define a mother to have been exposed to treatment if she indicates in the endline survey, that she received text messages on at least one of the relevant topics and correctly identifies PKH as the sender. Similar to approach 1 we present LATE estimates for specifications with and without controls.

$$\text{LATE} \qquad Exp_{ivt+1} = \eta + \theta T_i + \kappa' Y_{ivt} + \lambda' X_{ivt} + \mu' Z_{vt} + \phi_d + \nu_{iv} \tag{2a}$$

$$Y_{ivt+1} = \alpha + \beta \widehat{Exp}_{ivt+1} + \gamma' Y_{ivt} + \delta' X_{ivt} + \zeta' Z_{vt} + \phi_d + \epsilon_{iv} \tag{2b}$$

In all specifications standard errors are clustered at the village level. Moreover, to account for family-wise error rates we adjust the p-values using Westfall-Young stepdown adjusted p-values.[27]

4.2. Results

The main results are presented in Tables 3 (health behavior) and 4 (health outcomes). In each table Panel A refers to ITT estimates and Panel B to LATE estimands. For each dependent variable we always present results from two specifications: (i) unconditional effects estimating specifications 1 and 2 without controls and (ii) conditional effects including the control variables outlined above.[28]

[26] The following controls are included in X: age, education, subjective well-being, whether she is first-time mother, whether she is currently pregnant, the number of children living in the household, a personality measure of openness, a bargaining power index, attendance of PKH meetings, and the size of the social network among PKH mothers in the village. The following controls are included in Z: the availability of a midwife in the village, the number of early childhood facilities, the number of health centers, distance to the district capital.

[27] We define 'family' conservatively and include all health behavior, outcome, and knowledge variables (in total 13 variables). For implementation in Stata, we use the command `wyoung` (Jones et al., 2019).

[28] First-stage results from 2SLS related to our LATE specifications are presented in Table A.6 in Appendix A.3.

Table 3 about here.

Overall, we find that the sms campaign had a substantial positive impact on health behavior. Postnatal care practices improved by 8pp (or 15%), child vaccination rates by 7pp (9%), and hygiene practices by 10pp (50%).[29] All effects remain significant once including controls and once accounting for family-wise error rates. In comparison to our ITT results, LATE estimands show effect sizes that are about 2–3 pp higher.

Regarding health outcomes, we find that the sms capaign helped reduce maternal anemia rates. Mothers in the intervention group are about 10 pp (or 20%) less likely to have a moderate or severe form of anemia (Table 4). These effects seem to be driven primarily by older mothers (i.e. above 35), who in general face a higher likelihood of iron deficiencies (see Table B.20 in Online Appendix B.3).[30]

Turning towards child anthropometric outcomes our impact estimates for the main sample are statistically insignificant at conventional levels. Since our study design has insufficient statistical power to detect reasonable effect sizes for these outcomes, we discuss the economic relevance of the obtained estimates.[31] Similar to PKH's main impact evaluation (Cahyadi et al., 2020), we do not find any meaningful economic impact of the sms campaign on child wasting as coefficients are close to zero. In contrast, albeit statistically insignificant, we obtain suggestive evidence that the sms campaign might have reduced the incidence of child stunting by about 2.7pp to 3.4pp.

Since research has shown that young children might benefit from interventions very differently depending on their age (Attanasio et al., 2022; Cahyadi et al., 2020), we in addition study impacts for various age groups (Tables B.21 and B.22 in Appendix B). In fact, we find a marginally statistically significant effect of our intervention on the prevalence of stunting among children aged 19-36 months. For this age group the likelihood of stunting reduces by 8 pp, a decline of 21%. Bearing in mind that these children were between 4 and 22 months old at the start of the sms campaign, they might have benefited in particular from the improvement in postnatal care.[32]

[29] Please see Table B.19 in Online Appendix B.3 for additional results on child vaccinations. Defining vaccination practice as the number of vaccination the youngest child received, we show that the number of vaccinations increased by almost 90%).

[30] Looking more closely into impacts among mothers who were pregnant during the endline survey – mothers who were explicitly targeted in some of the anemia related text messages – we obtain large yet statistically insignificant treatment effects (column 4 in Table B.20); a result that might be attributable to a probable lack of statistical power due to the small sample size. Furthermore, among women who were pregnant during the intervention period, we find that self-reported health and iron pill intake might have increased as a result of the intervention (Table B.23 in Online Appendix B.3). The related coefficients show a relevant economic magnitude but , however, are statistically insignificant. This is likely due to the small sample size as power calculations suggest (see Table C.5 in the Online Appendix C)

[31] Impact evaluations from CCTs typically show small impacts in the range of 2pp—6pp (Attanasio et al., 2014; Cahyadi et al., 2020; Gertler, 2004). In comparison the minimum detectable effects in our study are large (6.5pp for wasting and 8.7pp for stunting as shown in Table C.3 in the online appendix.). Furthermore, given the measurement of short-term impacts and the light-touch nature of our intervention it seems less likely to expect sufficiently large impacts (in terms of achieving statistical significance) on child anthropometric outcomes.

[32] Child nutrition-related behavior was only captured for sub-samples of the mothers and we are likely

Table 4 about here.

4.3. Robustness

We conduct a number of robustness checks (see tables in Online Appendix B.2). First, we account for potential non-random attrition. As discussed in Section 3.3.3, there is some evidence for selective attrition. We assess whether our previous estimates are affected by differential attrition by first estimating treatment effect bounds following Lee (2009). As shown in Tables B.4 and B.5, the obtained treatment effects remain statistically and economically significant at the upper an lower bounds for all health practice outcomes as well as for the mother's anemia status. Following the strategy outlined in Kling and Liebman (2004), we derive alternative bounds varying the assumptions about health behavior & outcomes for respondents who could not be resurveyed at endline. We start with very extreme assumptions (i.e. setting the outcome values of non-interviewed respondents to the minimum value for mothers in the treatment arm and to the maximum value for mothers in the control arm) and then relax these assumptions step-by-step (see Tables B.6 and B.7). For health related behavior, effects hold throughout for hygiene practice and the overall practice index, they are relatively stable for vaccination practice; however, for postnatal care practices the estimated effect notably depends on the made assumptions. The same is true for the mother's anemia status. As an alternative approach, we weight the observations with the inverse probability of participating in the endline survey to obtain an estimation for population-wide effects and compare these to the estimated effects of our main specification. We predict endline survey participation with the individual and village characteristics that are described in Table A.1 as well as the outcome variables at baseline. Results are provided in Tables B.8 and B.9. All results hold. If at all, coefficient sizes slightly increase.

Second, we try to correct for potential spatial correlation in our error terms following Conley (1999).[33] As shown in Tables B.10 and B.11 all effects hold with the exception of the estimated effect on post-natal care practices when controls are not included. Third, we test whether results are sensitive to the original randomization process using randomization inference. All our effects hold (see Table B.12).[34] Fourth, we check whether our main results are sensitive to our preferred covariate specification. In Table

underpowered to estimate treatment effects. Information on breastfeeding practices was collected for mothers who gave birth during the intervention period (N=181), while child feeding practices was collected only for the sample of women who had children below the age of two at the time of the survey (N=535). The impact of the intervention on these outcomes are shown in Table B.24 in Online Appendix B.3. While almost all coefficients seem to indicate positive impacts of the sms campaign on breastfeeding outcomes with some coefficients (e.g. on exclusive breastfeeding) being of relevant economic magnitude, none of the coefficients are statistically significant. Therefore we consider this to be suggestive evidence only. Again, this is likely an issue of the small sample size (see Table C.5 in the Online Appendix). In terms of child feeding practices, we do not find any impact of the intervention on the provision of a healthy diet. Estimated effects are statistically insignificant and of no economic relevance.

[33] We implement the procedure in Stata using the `acreg` package (Colella et al., 2019).

[34] We implement Fisher's permutation-based randomization inference using `ritest` (Hess, 2017).

B.13, we present results from lasso methods in which covariates are selected based on their predictive power. All our previous results continue to hold.[35] Fifth, we include enumerator fixed effects (see Table B.14). All our effects hold, though they are slightly less precisely estimated for anemia status.

Sixth, we exclude all respondents who are not the biological mother but the father or another female relative who is the main care-taker of the child as they might differ in terms of pre- and post-natal behavior and health outcomes (in total 63 observations at endline). The estimated effects remain very similar both in size and level of significance (see Tables B.15 and B.16).

Seventh, we attempt to assess whether spillovers occurred. Assuming that spillovers are more likely to happen in a control village that is close to a treatment village compared to one that is further away we present in Table B.17 specifications in which control villages are classified according to the distance to the next treatment village.[36] We do not find any evidence for spillovers from treatment to control villages on health practices or health outcomes.

Finally, we analyze whether our intervention went hand-in-hand with a change in local health care supply, which could explain improvements in health practices. While the sms campaign did not involve the coordination with local health care providers, it could be that in treatment villages local health care providers increased supply to address increases in health care demand. To assess whether our intervention led and/or interacted with increases in local health care supply we leverage data from Indonesia's village census (PODES) from 2018 (i.e. two years after the intervention). Overall, we do not observe any differences in health care supply between treatment and control villages (see Table B.18). Assuming that any change in health care supply that was triggered by the intervention did not disappear after two years, we conclude that changes in local health care supply are unlikely to explain our findings.[37]

Overall, our findings are robust to a number of different specifications and sensitivity checks. All our main results hold with one exception: the estimated effect on postnatal care practices without controls loses significance when correcting standard errors for spatial correlation and is furthermore not robust when accounting for potential non-random attrition with Kling-Liebmann bounds. Our results on postnatal care practices need thus be interpreted with more care.

[35] As described in the related table note a large set of possible control variables was considered. For each estimation, the lasso procedure selected only very few controls (zero to six controls depending on the outcome variable).

[36] We define a village to be close if the distance to the next treatment village is below the median distance in the sample within a province (in both provinces, this is approx. 2.6 km).

[37] Health care supply at baseline (based on information from PODES 2014) did not differ between treatment and control villages (see Table A.3 in Appendix A.1). Since not all health supply indicators were already collected in PODES 2014, the regression specifications do not control for 'baseline' (PODES 2014) health supply indicators.

5. Mechanisms

We observe a substantial impact of the intervention on health related behavior and mother's anemia status. In the following, we attempt to shed light on what is driving these effects and what can explain the large variation in effect sizes that we observe.

5.1. Knowledge Gains, Reminders or Nudges

Examining the impact of the sms campaign on health knowledge (Table 5), we find that our intervention on average improved knowledge in three out of five health domains. The share of correct responses on postnatal care by 7 pp (36%), on hygiene by 3 pp (7%), and on anemia by 6 pp (or, 40%). Combined knowledge across all five domains improved by 4 pp (15%). In contrast, we find that the intervention on average did not improve mother's vaccination knowledge.

Table 5 about here.

To what extent the increase in health knowledge is responsible for the observed impact of the sms campaign on health behavior & outcomes is, however, not yet established. In the following, we analyze to what extent our findings are triggered through (i) the provision of new information (health knowledge gains), (ii) reminding mothers and help them recall previously stored information, and (iii) nudging mothers to adopt certain behaviors (de Vries et al., 2021). Our experimental set-up does not allow us to obtain causal estimates on the relative role of each of these factors.[38] Yet, we can derive testable predictions based on the following considerations.[39] First, if messages were primarily effective in changing health behavior & outcomes by providing new information, we should be able to observe knowledge gains in the respective health domains which in turn should then be correlated with respective improvements in health behavior & outcomes at the individual level. Second, if text messages rather worked as reminders, we would expect a positive correlation between baseline knowledge and behavioral change. Finally, if messages worked as nudges, we would not necessarily see any sort of strong correlation between behavioral change and knowledge gains or baseline knowledge.[40]

Figure 1 about here.

[38] The study design did not involve any cross-randomization. Mothers in the treatment group all receive identical messages at the same point in time. Moreover, each sms topic consisted of a mix of new information, reminders, and nudges.

[39] Note that as both, knowledge and behavior can be affected by the intervention simultaneously, the following analyses can only be interpreted as suggestive evidence.

[40] In contrast to the 'information' and 'reminder' channel, we consider the 'nudge channel' as a mechanisms through which individuals are steered towards a desirable health behavior & outcome. In this sense PKH mothers' do not necessarily need to have any knowledge about why something should be done. For example, mothers might decide to consume healthier food because they are asked to do so as part of our sms irrespective of whether they (i) know what anemia is and (ii) understand that healthier (iron-rich) food is helping to prevent anemia.

Following these predictions, we re-estimate our previous empirical specifications by including additional interaction effects. Figure 1 depicts the respective results by displaying coefficient on the aggregated impact (treatment effect + interaction effect) for health behavior (Panel A) and health outcomes (Panel B). In each of the two panels results in subfigure (a) relate to the knowledge channel and in subfigure (b) to the reminder channel.[41]

Overall, we find that channels appear to be highly domain and context specific. The information channel seems to matter in explaining observed improvements in hygiene practices and maternal anemia status as we observe a strong association between health knowledge gains in the respective dimension and subsequent improvements in behavior/outcomes. In contrast, the reminder channel seems to play a role in the context of child vaccinations as the impact of our intervention on child vaccinations seems to be positively associated with better baseline vaccination knowledge. Despite a positive impact of our intervention on postnatal knowledge, we find that improvements in postnatal behavior (keeping child health books) seem to be unrelated to baseline knowledge and knowledge gains. Consequently, we speculate that the nudge channel is an important factor in this context.

5.2. Barriers to Behavioral Change

Besides the role of information, reminders, and nudges, a number of other factors can help explain the success and failure of mHealth campaigns. Ultimately, altering a person's behavior is complex and depends on a number of individual and environmental factors. Guided by the framework of the Theory of Planned Behavior (TBP), we examine in this section which factors moderated the impact of the sms campaign on health behavior & outcomes.[42] The TPB is a psychological theory that links beliefs to intentions with the latter ones assumed to determine behavior. The TPB consists of three components: (i) attitude, (ii) perceived behavioral control, and (iii) subjective norms which together shape an individual's behavioral intentions (Ajzen, 1985; Fishbein and Ajzen, 1980). In the following, we discuss the empirical evidence on each of the three components separately. Results are shown in Table 6.

[41] Results in subfigure (a) are derived from specifications where the outcome variable is regressed on treatment status, knowledge growth and the interaction of the two, controlling for mother and village characteristics, the lagged dependent variable as well as baseline knowledge, while in subfigure (b) the outcome variable is regressed on treatment status, baseline knowledge and the interaction of the two, controlling for mother and village characteristics and the lagged dependent variable.

[42] Please see Brew-Sam and Chib (2020) for an overview and review of behavior change theories including the TPB and others such as social cognitive theory social learning theory, the elaboration likelihood model, protection motivation theory, the transtheoretical model of behavior change, the self-regulatory model, the integrative model, the health decision model, the theory of reasoned action, and coping theories. We selected the TPB as reference framework for two reasons. First, the theory has shown to have high predictive power for intentions in health-related areas such as condom use, leisure, exercise, and diets (Conner and Sparks, 2015; McEachan et al., 2011). Second, the TPB has a long tradition to be applied to a wide range of behaviors and cultural settings (Armitage and Conner, 2001; Appleby et al., 2015; Hsu et al., 2017; Protogerou et al., 2012).

Table 6 about here.

Attitude For developing the intention to change one's behavior it is conducive to have a favorable attitude towards the behavior in question. Since attitudes are often formed through learning and personal experiences, it is important that PKH women can (i) overcome status quo bias and (ii) be willing to engage in new practices and behaviors.

In the following, we implement two different empirical tests to gauge the role of attitudes in affecting our main results. The first one relates more closely to status quo bias and is based on the assumption that first-time mothers are more likely to follow the recommendations from the mHealth BCC intervention, since certain habits and behavioral patterns might not yet have manifested themselves. The second one relates specifically to a person's willingness to engage in new practices and is closely linked to existing research on character traits. More specifically, a person's personality in terms of openness to experience has been linked to predicting attitude and behavioral change with respect to technology adoption and health-related practices (Bano et al., 2019; Eisenman et al., 1980; Ramírez-Correa et al., 2019).

Results are shown in Panel A and Panel B of Table 6. Overall, we find some supportive evidence that a mother's attitude towards the intervention plays a role for behavioral change. First time mothers and mothers who are more open in terms of their personality are more likely to improve their postnatal care and their vaccination practices. Also, the positive impact of the intervention on reducing anemia rates is mostly driven by first-time mothers.

Perceived behavioral control For a change in health behavior & outcomes to occur, a person must belief that the behavior in question is under his or her control. Measurement of a person's perceived behavioral control typically relies on self-reported instruments with a focus on two parts (i) self-efficacy and (ii) controllability (Ajzen, 2002). While our data does not allow to examine the role of self-efficacy, we are able to shed light on the importance of controllability, the extent to which a behavior is up to the individual, as a moderating factor. Proxying controllability by a bargaining index that is derived by principal component analysis from mothers' perceived decision-making power on a number of child and household related dimensions, we show in Panel C of Table 6 the related empirical results. Bearing in mind limitations in terms of measurement, we overall find little evidence that mothers' perceptions about their behavioral control affected health behavior (postnatal care, vaccinations, hygience practices) and outcomes (maternal anemia, a child's wasting status). However, we see a differential effect on stunting. Treated mothers with higher bargaining power seem to be less likely to have a child that is stunted.

Subjective norms An individual's perception about prevalent normative beliefs and subjective norms regarding the health behavior in question can influence a person's intentions to embrace a different health practice. In this context social pressure and judgements of others might matter for health behavior change. In Panel D of Table 6,

we present results from specifications that proxy prevailing norms using a village-level classification of baseline health practices.[43] Overall, we find little differential effects, with the exception of anemia rates.

Summary Exploring our findings through the lens of the TPB and bearing in mind measurement constraints, we find in general little support for the importance of perceived behavioral control factors and locally prevalent subjective norms in moderating the impact of our intervention. In contrast, we find some supporting evidence for the view that a favorable attitude towards the sms campaign helped improve some health behaviors & outcomes (postnatal care practices, child vaccinations, and maternal anemia rates).

5.3. Supply-Side Constraints

Considering that a large body of research in public health has stressed the importance of local-level health care availability in influencing health outcomes and the success of health interventions (Gage et al., 2016; Kemp et al., 2018; Verma and Dash, 2021), we examine whether the impact of our sms campaign is related to supply-side issues. Proxying local-level supply-side infrastructure by the availability of (i) the density of health centers, (ii) the density of early childhood facilities, and (iii) the availability of a village midwife, we overall find that differences in local-level health infrastructure seem not to mediate the impact of our intervention (Table 7).

Table 7 about here.

6. Discussion

The evidence so far suggests that the sms campaign was successful in improving a number of maternal health indicators. This raises several key questions: Why did the mHealth campaign succeed (at least to a certain extent), while PKH's cash benefits and group meetings were not able to achieve sizeable gains in health knowledge, behavior, and outcomes in the short-run (Cahyadi et al., 2020)? Assuming that the intervention was successful, is its implementation advisable taking into account its costs?

6.1. On the Importance of PKH Groups

According to official implementation guidelines the monthly PKH group meetings are meant to be an important means to improve health outcomes. Led by a trained facilitator participation in these groups is expected to (i) generate health knowledge among PKH mothers and (ii) nudge PKH mothers to comply with beneficial health practices (health conditions and its related monitoring + financial sanctions). Field evidence, however,

[43] From the baseline data we calculate the share of women following good health practices. We then construct a dummy variable that takes the value 1 if a village ranks above the median in terms of good health practices and 0 otherwise.

has shown that the de-facto implementation of PKH groups suffers from a number of problems that undermine PKH's health objectives. First, PKH groups often do not reconvene on a monthly basis; and even if PKH group meetings take place, only a fraction of PKH beneficiaries tend to attend these meetings on a monthly basis (about 30 percent of women in our sample). Second, the vast majority of the time during the PKH meetings is ultimately devoted to discussions focusing on PKH cash disbursements. Therefore, little time is usually spent discussing health related topics (C4ED, 2019). Third, conditionality criteria are often not monitored and sanctions are rare (C4ED, 2019).

Given these shortcoming and taking into account that PKH cash benefits are rather moderate compared to other CCTs such as Bolsa Familia and Progresa (Oportunidates/ Prospera), it is perhaps less surprising that PKH's main impact evaluation had found little impacts of the program on health knowledge and behaviors (Cahyadi et al., 2020). Leveraging our own data from this study, we find supporting evidence for the view that PKH group meetings seem to contribute little to improvements in health indicators. As shown in Table 8 we find that neither the (regular) attendance of PKH meetings nor a PKH mother's social PKH network seem to mediate the impact of our intervention. Yet, it seems that our sms campaign was able to at least partially address some of PKH's implementation gaps. Besides creating knowledge on and reminding PKH mothers of certain health topics, the sms campaign managed to nudge mothers into adopting beneficial health practices.

Table 8 about here.

6.2. Cost Effectiveness

Lastly, we want to discuss the cost-effectiveness of our intervention. In the following we put our results into perspective with respect to (i) CCTs and (ii) mHealth interventions. Considering the absence of systematic reviews and meta-studies on the cost effectiveness of health outcomes in CCTs and mHealth interventions and taking into account that intervention costs are often not tabulated (Hall et al., 2015; Hazel et al., 2021; Lagarde et al., 2009), we compiled our own estimates based on a review of the literature.[44] In the following we define cost-effectiveness as the impact achieved (in percentage points) divided by an intervention's cost (direct and indirect costs) per household in 2018 US dollars. We refer to this measure as the IC ratio.

Before we compare IC ratios, we turn to the average costs of our intervention. At the time of implementation direct costs for sending 156 sms amounted to 0.50 US Dollars per household, while indirect costs (administrative, technical equipment) amounted to 1 US Dollar per household. In 2018 prices the intervention costed about 1.66 US Dollar per household. Column 10 in Panel B of Table 9 depicts IC ratios for our main health outcomes.

[44] It should be noted that our estimates are sometimes not strictly comparable since studies have used related but different outcome variables and alternative ways in defining outcomes.

Table 9 about here.

Panel A of Table 9 provides estimates of the IC ratio for different CCTs.[45] While most CCTs seem to achieve positive health impacts across a broad set of health indicators, we find IC ratios ranging from 0.002 to 0.022 for the major government-implemented CCTs (Bolsa Familia, Familias en Accion, PKH, Progresa) and higher IC ratios for the exclusively health-focused CCTs in Nigeria and Kenya. In comparison, IC estimates from our sms intervention (Panel B) appear much larger. Though the type of health indicators evaluated differ strongly across studies, the most comparable indicator relates to child vaccinations. With respect to child vaccinations our sms intervention compares well against CCT programs.[46]

Next, we compare our intervention to the IC ratios from other mHealth studies (Panel B in Table 9); a comparison that is more straightforward given the nature of our intervention. Bearing in mind that we were not able find studies that use exactly the same outcome variable (definition), we find that our intervention compares reasonably well against other mHealth studies, in particular since mHealth studies that report costs appear to be biased towards studies that tend to show positive intervention results (Brown and Skelly, 2019; Hazel et al., 2021).

Lastly, given concerns about selective reporting of costs in mHealth studies and considering that mHealth sms interventions are rather inexpensive, we abstract from cost considerations and focus on simply comparing the impact of our intervention (column 6 in Table 9) to the reported impact of other mHealth studies (Table C.6 in the Online Appendix C). Bearing in mind large variations in the intervention design and outcome measurement, our impacts appear to be in the normal range of reported impacts of sms campaigns. In summary, we find that the our sms campaign compares rather well against other published studies in terms of impacts and costs.

7. Conclusion

In this paper, we examined to what extent a light-touch and inexpensive mHealth BCC intervention can be an effective tool in improving maternal and child health behavior & outcomes among recipients of cash transfer programs. To address this question, we study a clustered RCT implemented between 2014 and 2016 involving 127 villages and 1,821 women with young children in Indonesia.

[45] The CCTs we selected fall into two groups. All selected CCTs focus on young families and involve health conditionality criteria. Besides PKH three other CCTs (Bolsa Familia, Familias en Accion, Progresa (Opportunidades/Prospera) involve education related conditionality criteria.

[46] It should be noted that estimates of program costs for CCTs include costs that are not related to health. For example, some CCTs provide households with cash to send their children to school. Therefore, our cost estimates are likely to overstate CCT implementation costs with respect to health outcomes. However, even under some moderate assumptions, e.g. that only 1 percent of a CCT's implementation costs are devoted to health outcomes, the IC ratio of our intervention compares well. Please see Section C.4 in the Online Appendix for a detailed overview on how estimates were compiled.

Our study finds strong improvements in maternal health behavior and outcomes as a result of the mHealth BCC campaign. Health behavior improved in all three targeted indicators (postnatal care, child vaccinations, hygiene practices), while the prevalence of maternal anemia was reduced. In addition, our study provides suggestive evidence that the sms campaign might have improved certain child anthropometric outcomes. While the intervention certainly had no impact on the incidence of wasting among children below the age of five, we obtain some evidence that the intervention contributed to lowering child stunting rates.

Furthermore, our research sheds light on mediating factors that help explain the main results. First, we show that the sms campaign achieved sizeable health knowledge gains among PKH mothers exposed to our intervention. Mothers became substantially more knowledgeable with respect to three out of five topics (anemia, hygiene and postnatal care) the exception being knowledge on breastfeeding and vaccination. To what extent these gains in health knowledge are responsible for the observed improvements in health behavior & outcomes is debatable and might depend on the specific health dimension under consideration. We find that knowledge gains are positively correlated with improvements in mothers' hygiene practices and anemia reduction but unrelated to the found improvements in postnatal care behavior and child vaccinations. Regarding these two latter indicators we believe that other campaign features play an important role (reminder and nudge function of the sms).

Second and adopting a behavior change theory framework, we show that favorable attitudes by mothers towards the intervention (proxied through indicators for status quo bias and personality traits (openness)) seems to have at least partially facilitated the impact of our intervention. Lastly, we demonstrate that the impact of our intervention on health behavior & outcomes is not affected by local-level health care supply side constraints.

Overall, we believe the results demonstrate that mHealth BCC interventions can be a cost-effective tool to improve crucial maternal and child health indicators among recipients of cash transfer programs. Our moderation analysis, however, points out that effects can be very context-specific. In fact, our mHealth BCC sms-campaign appears to suggest that main channels and mechanisms differ across various health dimensions and indicators.

In closing, we would like to point out that there are two important limitations of the intervention that could affect the transferability of our findings to other contexts. First, our study sample comprises CCT beneficiaries which might differ from the general population and from poor non-CCT beneficiaries in several ways. Second, our results provide evidence for short-term impacts only. Due to the lack of a longer-term data collection process, we do not know whether improvements in mothers' health knowledge, behavior, and outcomes (anemia) persisted in the medium- and long run.

References

Agravat, N. (2013). *mHealth and neonatal resuscitation: A review of interventions, approaches and applications*. mHealth Alliance.

Ahmed, A., M. Hidrobo, J. Hoddinott, B. Koch, S. Roy, and S. Tauseef (2020). Social protection and sustainable poverty reduction: Experimental evidence from bangladesh. *IFPRI Discussion paper 01988*.

Ahmed, A., J. Hoddinott, and S. Roy (2019). Food transfers, cash transfers, behavior change communication and child nutrition. *IFPRI Discussion paper 01868*.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. Action-control: From cognition to behavior, pp. 11–39. Springer.

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior1. *Journal of Applied Social Psychology 32*(4), 665–683.

Alatas, V., A. Banerjee, R. Hanna, B. Olken, R. Purnamasari, and M. Wai-Poi (2019). Does elite capture matter? local elites and targeted welfare programs in indonesia. *American Economic Review Papers & Proceedings 109*, 334–339.

Appleby, B., C. Roskell, and W. Daly (2015). What are health professionals' intentions toward using research and products of research in clinical practice? a systematic review and narrative synthesis. *Nursing open 3*(3), 125–139.

Armanasco, A., Y. Miller, B. Fjeldsoe, and A. Marshall (2017). Preventive health behavior change text message interventions: A meta-analysis. *American Journal of Preventive Medicine 52*(3), 391–402.

Armitage, C. and M. Conner (2001). Efficacy of the theory of planned behaviour: a meta-analytic review. *The British journal of social psychology 40*(4), 471–499.

Arriagada, A., L. Rawlings, J. Trias, and M. Zumaeta-Aurazo (2020). Combining cash transfers and parenting programs: Operational design considerations. *Global Social Policy 20*(1), 15–20.

Athey, S. and G. Imbens (2017). The econometrics of randomized experiments. Volume 1 of *Handbook of Economic Field Experiments*, pp. 73–140. North-Holland.

Attanasio, O., H. Baker-Henningham, R. Bernal, C. Meghir, D. Pineda, and M. Rubio-Codina (2022). Early Stimulation and Nutrition: The Impacts of a Scalable Intervention. *Journal of the European Economic Association*.

Attanasio, O., C. Fernandez, E. Fitzsimons, S. Grantham-McGregor, C. Meghir, and M. Rubio-Codina (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *BMJ 349*.

Attanasio, O., L. Gómez, P. Heredia, and M. Vera-Hernández (2014). The short-term impact of a conditional cash subsidy on child health and nutrition in colombia. *Report of the Institute for Fiscal Studies*.

Avula, R., P. Menon, K. Saha, M. Bhuiyan, A. Chowdhury, S. Siraj, R. Haque, C. Jalal, K. Afsana, and E. Frongillo (2013). A program impact pathway analysis identifies critical steps in the implementation and utilization of a behavior change communication intervention promoting infant and child feeding practices in bangladesh. *The Journal of nutrition 143*.

Badawy, S., M. Sherif, and L. Kuhns (2017). Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: A systematic review. *JMIR Mhealth Uhealth 5*(4), e50.

Bahety, G., S. Bauhoff, D. Patel, and J. Potter (2021). Texts don't nudge: An adaptive trial to prevent the spread of covid-19 in india. *Journal of Development Economics 153*, 102747.

Baird, S., F. Ferreira, B. Özler, and M. Woolcock (2014). Conditional, unconditional and everything in between: A systematic review of the effects of cash transfer programs on schooling outcomes. *Journal of Development Effectiveness 6*(1), 1–43.

Baird, S., G. McIntosh, and B. Özler (2011). Cash or condition? evidence from a cash transfer experiment. *Quarterly Journal of Economics 126*(4), 1709–1753.

Banerjee, A., M. Alsan, E. Breza, A. Chandrasekhar, A. Chowdhury, P. Goldsmith-Pinkham, and B. Olken (2020). Messages on covid-19 prevention in india increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities. *NBER working paper no. 27496*.

Banerjee, A., D. Karlan, R. Osei, H. Trachtman, and C. Udry (2021). Unpacking a multifaceted program to build sustainable income for the very poor. *Journal of Development Economics*, 102781.

Bano, S., U. Shah, and S. Ali (2019). Personality and technology: Big five personality traits as descriptors of universal acceptance and usage of technology utaut. *University of Nebraska working papers*.

Baron-Cohen, S. (2000). *Theory of mind and autism: A fifteen-year review*. MIT Press.

Bastagli, F., J. Hagen-Zanker, L. Harman, V. Barca, G. Sturge, T. Schmidt, and L. Pellerano (2016). *Cash transfers: what does the evidence say? A rigorous review of programme impact and of the role of design and implementation features*. Overseas Development Institute.

Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen (2015). Turning a shove into a nudge? a "labeled cash transfer" for education. *American Economic Journal: Economic Policy 7*(3), 86–125.

Bhutta, Z., J. Das, A. Rizvi, M. Gaffey, N. Walker, S. Horton, P. Webb, A. Lartey, and R. Black (2013). Evidence-based interventions for improvement of maternal and child nutrition: What can be done and what cost? *The Lancet 382*(9890), 452–477.

Biscaye, P., D. Coomes, E. Ebeling, A. Favreau, and T. Madsen (2017). Review of evidence on long-term impacts of cash transfer programs. *mimeo*.

Blaya, J., H. Fraser, and B. Holt (2010). E-health technologies show promise in developing countries. *Health Affairs 29*(2), 244–251.

Brew-Sam, N. and A. Chib (2020). *Theoretical advances in mobile health communication research: An Empowerment Approach to Self-Management*. Elsevier.

Brollo, F., K. Kaufmann, and E. L. Farrara (2017). Learning about the enforcement of conditional welfare programs: Evidence from brazil. *IZA working paper no. 10654*.

Brown, A. and H. Skelly (2019). How much evidence is there really? mapping the evidence base for ict4d interventions. *Information Technologies and International Development*.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

Burn, E., S. Nghiem, S. Jan, J. Redfern, A. Rodgers, A. Thiagalingam, N. Graves, and C. Chow (2017). Cost-effectiveness of a text message programme for the prevention of recurrent cardiovascular events. *Heart 103*(12), 893–894.

C4ED (2019). Implementation research report: Assessment of pkh implementation and expansion in indonesia. *Center for Evaluation and Development*.

Cahyadi, N., R. Hanna, B. Olken, R. Prima, E. Satriawan, and E. Syamsulhakim (2020). Cumulative impacts of conditional cash transfer programs: Experimental evidence from indonesia. *American Economic Journal: Economic Policy 12*(4), 88–110.

Carneiro, P., L. Kraftman, G. Mason, L. Moore, I. Rasul, and M. Scott (2021). The impacts of a multifaceted prenatal intervention on human capital accumulation in early life. *American Economic Review 111*(8), 2506–49.

Celik, R. and E. Toruner (2020). The effect of technology-based programmes on changing health behaviours of adolescents: Systematic review. *Comprehensive Child and Adolescent Nursing 43*(2), 92–110.

Christian, C., L. Hensel, and C. Roth (2019). Income shocks and suicide: Causal evidence from indonesia. *Review of Economics and Statistics 101*(5), 905–920.

Cole-Lewis, H. and T. Kershaw (2010). Text messaging as a tool for behaviour change in disease prvention and management. *Epidemiological Review 32*(1), 55–69.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). Inference with arbitrary clustering. *IZA working paper no. 12584*.

Conley, T. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–459.

Conner, M. and P. Sparks (2015). Theory of planned behaviour and the reasoned action approach. Predicting and changing health behaviour: research and practice with social cognition models, pp. 142–188. Open University Press.

Cortes, K., H. Fricke, S. Loeb, D. Song, and B. York (2020). Too little or too much? actionable advise in an early-childhood text messaging experiment. *Education Finance and Policy 16*(1), 1–44.

Crutzen, R., J. de Nooijer, W. Brouwer, A. Oenema, J. Brug, and N. de Vries (2011). Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: A systematic review. *Health Education & Behavior 38*(1), 49–62.

Dai, H., S. Saccardo, and M. Han (2021). Behavioural nudges increase covid-19 vaccinations. *Nature 597*, 404–409.

Dammert, A., J. Galdo, and V. Galdo (2014). Preventing dengue through mobile phones: Evidence from a field experiment in peru. *Journal of Health Economics 35*(May), 147–161.

Datta, S., S. Sekhar, P. Ranganathan, and K. Sivakumar (2014). A study to assess the feasibility of text messaging service in delivering maternal and child healthcare messages in a rural area of tamil nadu, india. *The Australasian medical journal 7*, 175–80.

de Vries, M., M. Rieger, R. Sparrow, E. Prafiantini, and R. Agustina (2021). Snacks, nudges and asymmetric peer influence: Evidence from food choice experiments with children in indonesia. *Journal of Health Economics 79*, 102508.

DHS (2018). *Indonesia: Demographic and Health Survey 2017*. National Population and Family Planning Board, Statistics Indonesia, Ministry of Health and DHS Program.

Doss, C., E. Fahle, S. Loeb, and B. York (2019). More than just a nudge: Supporting kindergarten parents with differentiated and personalized text messages. *American Economic Journal: Applied Economics 54*(3), 567–603.

Eisenman, R., J. Grossman, and R. Goldstein (1980). Undergraduate marijuana use as related to internal sensation novelty seeking and openness to experience. *Journal of Clinical Psychology 36*(4), 1013–1019.

Eze, P., L. Lawani, and Y. Acharya (2021). Short message service (sms) reminders for childhood immunisation in low-income and middle-income countries: a systematic review and meta-analysis. *BMJ Global Health 6*(7).

Falco, P. and S. Zaccagni (2021). Promoting social distancing in a pandemic: Beyond good intentions. *PLOS ONE 16*(12).

Fetter, D., M. Dharmar, S. Lawry-Hall, J. Pressman, J. Chapman, and R. Scherr (2019). The influence of gain-framed and loss-framed health messages on nutrition and physical activity knowledge. *Global Pediatric Health 6*, 2333794X19857405.

Fiala, N. (2018). Returns to microcredit, cash grants and training for male and female microentrepreneurs in uganda. *World Development 105*(C), 189–200.

Field, E. and E. Maffioli (2021). Are behavioral change interventions needed to make cash transfer programs work for children? experimental evidence from myanmar. *NBER working paper no. 28443*.

Filmer, D. and N. Schady (2009). Are there diminishing returns to transfer size in conditional cash transfers? *World Bank policy research working paper no. 4999*.

Fishbein, M. and I. Ajzen (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs.

Fiszbein, A., N. Schady, F. Ferreira, M. Grosh, N. Kelleher, P. Olinto, and E. Skoufias (2009). *Conditional cash transfers: Reducing present and future poverty*. World Bank.

Fjelsoe, B., A. Marshall, and Y. Miller (2009). Behavior change interventions delivered by mobile phone short message service. *American Journal of Preventive Medicine 36*(2), 165–173.

Fonagy, P., M. Steele, H. Steele, G. Moran, and A. Higgitt (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal 12*(3), 201–218.

Fricke, H., D. Kalogrides, and S. Loeb (2018). It's too annoying: Who drops out of educational text messaging programs and why. *Economics Letters 173*, 39–43.

Gaarder, M., A. Glassman, and J. Todd (2010). Conditional cash transfers and health: unpacking the causal chain. *Journal of Development Effectiveness 2*(1), 6–50.

Gage, A., O. Ilombu, and A. Akinyemi (2016). Service readiness, health facility management practices, and delivery care utilization in five states of nigeria: a cross-sectional analysis. *BMC Pregnancy Childbirth 16*(297).

Gallegos, D., R. Russell-Bennett, J. Previte, and J. Parkinson (2014). Can a text message a week improve breastfeeding? *BMC pregnancy and childbirth 14*, 374.

Gertler, P. (2004). Do conditional cash transfers improve child health? evidence from progresa's control randomized experiment. *American Economic Review Papers & Proceedings 94*, 336–341.

Gilligan, D., A. Arrieta, S. Devereux, J. Hoddinott, D. Kebede, N. Ledlie, K. Roelen, and A. Taffesse (2020). Impact evaluation of improved nutrition through integrated basic social services and social cash transfer pilot program (in-sct) in oromia and snnp regions, ethiopia. *Unicef*.

Glassman, A., D. Duran, L. Fleisher, D. Singer, R. Sturke, G. Angeles, J. Charles, B. Emrey, J. Gleason, W. Mwebsa, K. Saldana, K. Yarrow, and M. Koblinsky (2013). Impact of conditional cash transfers on maternal and newborn health. *Journal of Health, Population and Nutrition 31*(4), 48–66.

Grepin, K., J. Habyarimana, and W. Jack (2019). Cash on delivery: Results of a randomized experiment to promote maternal health care in kenya. *Journal of Health Economics 65*, 15–30.

Guanais, F. (2013). The combined effects of the expansion of primary health care and conditional cash transfers on infant mortality in brazil, 1998–2010. *American Journal of Public Health 103*(11), 2000–2006.

Gurman, T., S. Rubin, and A. Roess (2012). Effectiveness of mhealth behavior change communication interventions in developing countries: A systematic review of the literature. *Journal of health communication 17 Suppl 1*, 82–104.

Hall, A., H. Cole-Lewis, and J. Bernhardt (2015). Mobile text messaging for health: A systematic review of reviews. *Annual Review of Pulbic Health 36*, 393–415.

Haushofer, J. and J. Shapiro (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *Quarterly Journal of Economics 131*(4), 1973–2042.

Hazel, C., S. Bull, E. Greenwell, M. Bunik, J. Puma, and M. Perraillon (2021). Systematic review of cost-effectiveness analysis of behavior change communication apps: Assessment of key methods. *Digital Health 7*, 20552076211000559.

Hess, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal 17*(3), 630–651.

Hsu, C.-L., C.-Y. Chang, and C. Yansritaku (2017). Exploring purchase intention of green skincare products using the theory of planned behavior: Testing the moderating effects of country of origin and price sensitivity. *Journal of Retailing and Consumer Services 34*, 145–152.

Jiang, H., M. Li, L. Wen, Q. Hu, D. Yang, G. He, L. Baur, M. Dibley, and X. Qian (2014). Effect of short message service on infant feeding practice: findings from a community-based study in Shanghai, China. *JAMA pediatrics 168*(5), 471–478.

Jones, D., D. Molitor, and J. Reif (2019). What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study*. *The Quarterly Journal of Economics 134*(4), 1747–1791.

Karlan, D., R. Knight, and C. Udry (2015). Consulting and capital experiments with microenterprise tailors in ghana. *Journal of Economic Behavior & Organization 118*, 281–302.

Kemp, C., R. Sorensen, N. Puttkammer, R. Pierre, J. Honoré, L. Lipira, and C. Adolph (2018). Health facility readiness and facility-based birth in haiti: a maximum likelihood approach to linking household and facility data. *Journal of global health reports 2*.

Khan, G., S. Kureishy, N. Akbar, M. Nasir, M. Hussain, and I. Ahmed (2019). A stunting prevention cluster randomized controlled trial: Leveraging the social protection system to prevent stunting in district rahim yar khan, punjab, pakistan. Technical report, World Food Programme.

Khorshid, M., P. Afshari, and P. Abedi (2014). The effect of sms messaging on the compliance with iron supplementation among pregnant women in iran: a randomized controlled trial. *Journal of telemedicine and telecare 20*(4), 201–206.

Kling, J. R. and J. B. Liebman (2004, March). Experimental Analysis of Neighborhood Effects on Youth. Working Papers 1, Princeton University, Department of Economics, Industrial Relations Section.

Kusuma, D., J. Cohen, M. McConnel, and P. Berman (2016). Can cash transfers improve determinants of maternal mortality? evidence from the household and community programs in indonesia. *Social Science & Medicine 163*, 10–20.

Ladhani, S. and K. Sitter (2020). Conditional cash transfers: A critical review. *Development Policy Review 38*, 28–41.

Lagarde, M., A. Haines, and N. Palmer (2009). The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. *The Cochrane database of systematic reviews 4*(4).

Langendorf, C., T. T. Roederer, S. de Pee, D. Brown, S. Doyon, A.-A. Mamaty, L. Toure, W. Lynda, M. Manzo, and R. Grais (2014). Preventing acute malnutrition among young children in crises: A prospective intervention study in niger. *PLOS Medicine 11*(9), 1–15.

Lau, Y., T. Cassidy, D. Hacking, K. Brittain, H. Haricharan, and M. Heap (2014). Antenatal health promotion via short message service at a midwife obstetrics unit in south africa: A mixed methods study. *BMC pregnancy and childbirth 14*, 284.

Lee, D. S. (2009, 07). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies 76*(3), 1071–1102.

Levere, M., G. Acharya, and P. Bharadwaj (2016). The role of information and cash transfers on early childhood development : evidence from Nepal. Policy Research Working Paper Series 7931, The World Bank.

Little, M., K. Roelen, B. Lange, J. Steinert, A. Yakubovich, L. Cluver, and D. Humphreys (2021). Effectiveness of cash-plus programmes on early childhood outcomes compared to cash transfers alone: A systematic review and meta-analysis in low- and middle-income countries. *PLOS Medicine 18*(9), 1–24.

Loescher, L., S. Rains, S. Kramer, C. Akers, and R. Moussa (2018). A systematic review of interventions to enhance healthy lifestyle behaviors in adolescents delivered via mobile phone text messaging. *American Journal of Health Promotion 32*(4), 865–879.

Lund, S., B. Nielsen, M. Hemed, I. Boas, A. Said, K. Said, M. Makungu, and V. Rasch (2014). Mobile phones improve antenatal care attendance in zanzibar: A cluster randomized controlled trial. *BMC pregnancy and childbirth 14*, 29.

Majeed-Ariss, R., E. Baildam, M. Campbell, A. Chieng, D. Fallon, A. Hall, J. McDonagh, S. Stones, W. Thomson, and V. Swallow (2015). Apps and adolescents: A systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. *J Med Internet Res 17*(12), e287.

Mani, A., S. Mullainathan, S. Shafir, and J. Zhao (2013). Poverty impedes cognitive functioning. *Science 341*(6149), 976–980.

McEachan, R., M. Conner, N. Taylor, and R. Lawton (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: a meta-analysis. *Health Psychology Review 5*(2), 97–144.

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

MoSA (2020). *What is program keluarga harapan?* Ministry of Social Affairs.

MSC (2019). *Report on operations assessment and impact evaluation of Program Keluarga Harapan (PKH)*. MicroSave Consulting.

Nazara, S. and S. Rahayu (2013). Program keluarga harapan (pkh): Indonesian conditional cash transfer programme. *International Policy Centre for Inclusive Growth research brief no. 42*.

Okeke, E. and I. Abubakar (2020). Healthcare at the beginning of life and child survival: Evidence from a cash transfer experiment in nigeria. *Journal of Development Economics 143*(102426).

Orr, J. and R. King (2015). Mobile phone sms messages can enhance healthy behaviour: A meta-analysis of randomized controlled trials. *Health Psychology Review 9*(4), 397–416.

Palmer, M., J. Sutherland, S. Barnard, A. Wynne, E. Rezel, A. Doel, L. Grigsby-Duffy, S. Edwards, S. Russell, E. Hotopf, P. Perel, and C. Free (2018). The effectiveness of smoking cessation, physical activity/diet and alcohol reduction interventions delivered by mobile phones for the prevention of non-communicable diseases. *PLOS ONE 13*(1), 1–71.

Palmer, T., N. Batura, J. Skordis, O. Stirrup, F. Vanhuyse, and A. Copas (2022). Economic evaluation of a conditional cash transfer to retain women in the continuum of care during pregnancy, birth and the postnatal period in kenya. *PLOS Global Public Health 2*(3).

Parker, S. and P. Todd (2017). Conditional cash transfers: The case of 'progresa/oportunidades'. *Journal of Economic Literature 55*(3), 866–915.

Payne, H., C. Lister, J. West, and J. Bernhardt (2015). Behavioral functionality of mobile apps in health interventions: A systematic review of the literature. *JMIR mHealth uHealth 3*(1), e20.

Poorman, E., J. Gazmararian, R. Parker, B. Yang, and L. Elon (2014). Use of text messaging for maternal and infant health: A systematic review of the literature. *Maternal and child health journal 19*.

Pop-Eleches, C., H. Thirumurthy, J. Habyarimana, J. Zivin, M. Goldstein, D. de Walque, L. MacKeen, J. Haberer, S. Kimaiyo, J. Sidle, D. Ngare, and D. Bangsberg (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.

Premand, P. and O. Barry (2022). Behavioral change promotion, cash transfers and early childhood development: Experimental evidence from a government program in a low-income setting. *Journal of Development Economics*.

Protogerou, C., A. Flisher, L. Aaro, and C. Mathews (2012). The theory of planned behaviour as a framework for predicting sexual risk behaviour in sub-saharan african youth: a critical review. *Journal of Child & Adolescent Mental Health 24*, 15–35.

Raifman, J., H. Lanthorn, S. Rokicki, and G. Fink (2014). The impact of text message reminders on adherence to antimalarial treatment in northern ghana: A randomized trial. *PLoS ONE 9*(10).

Ramírez-Correa, P., E. Grandon, J. Alfaro-Pérez, and G. Painén-Aravena (2019). Personality types as moderators of the acceptance of information technologies in organizations: A multi-group analysis in pls-sem. *Sustainability*, 3987.

Rivera, J., D. Sotres-Alvarez, J. Habicht, T. Shamah, and S. Villalpando (2004). Impact of the mexican program for education, health, and nutrition (progresa) on rates of growth and anemia in infants and young children. *Journal of the American Medical Association 291*, 2563–2570.

Roelen, K., S. Devereux, Abdulai, B. Martorano, T. Palermo, and L. Ragno (2017). How to make 'cash plus' work: Linking cash transfers to services and sectors. *Innocenti Working paper no. 2017-10*.

Roy, S., M. Hidrobo, J. Hoddinott, and A. Ahmed (2019). Transfers, Behavior Change Communication, and Intimate Partner Violence: Postprogram Evidence from Rural Bangladesh. *The Review of Economics and Statistics 101*(5), 865–877.

Schilbach, F., H. Schofield, and S. Mullainathan (2016). The psychological lives of the poor. *American Economic Review: Papers & Proceedings 106*(5), 435–440.

Sedlmayr, R., A. Shah, and M. Sulaiman (2020). Cash-plus: Poverty impacts of alternative transfer-based approaches. *Journal of Development Economics 144*, 102418.

Segura-Pérez, S., R. Grajeda, and R. Pérez-Escamilla (2016). Conditional cash transfer programs and the health and nutrition of latin american children. *PMID 40*(2), 124–137.

Shei, A., F. Costa, M. Reis, and A. Koz (2014). The impact of brazil's bolsa família conditional cash transfer program on children's health care utilization and health outcomes. *BMC International Health and Human Rights 14*(10).

Shiferaw, S., M. Spigt, M. Tekie, M. Abdullah, M. Fantahun, and G. Dinant (2016). The effects of a locally developed mhealth intervention on delivery and postnatal care utilization: A prospective controlled evaluation among health centres in ethiopia. *PLoS ONE 11*(7), e0158600.

Slade, A. (2008). *The implications of attachment theory and research for adult psychotherapy: Research and clinical perspectives*. The Guilford Press.

Sondaal, S., J. Browne, M. Amoakoh-Coleman, A. Borgstein, A. Miltenburg, M. Verwijs, and K. Klipstein-Grobusch (2016). Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: A systematic review. *PLoS One 11*(5).

Tauchmann, H. (2012, June). LEEBOUNDS: Stata module for estimating Lee (2009) treatment effect bounds. Statistical Software Components, Boston College Department of Economics.

Thaler, R. (2018). From cashew to nudges: The evolution of behavioral economics. *American Economic Review 108*(6), 1265–1287.

Tirivayi, N., M. Knowles, and B. Davis (2016). The interaction between social protection and agriculture: A review of evidence. *Global Food Security 10*, 52–62.

Triyana, M. (2016). Do health care providers respond to demand-side incentives? evidence from indonesia. *American Economic Journal: Economic Policy 8*(4), 255–288.

Unicef (2018). *Designing digital interventions for lasting impact: A human-centred guide to digital health deployment*. United Nations Children's Fund.

Urquieta, J., G. Angeles, T. Mroz, H. Lamdrid-Figueroa, and B. Hernández (2009). Impact of oportunidades on skilled attendance at delivery in rural areas. *Economic Development and Cultural Change 57*(3), 539–558.

Verma, V. and U. Dash (2021). Supply-side readiness for universal health coverage: Assessing service availability and barriers in remote and fragile setting. *Journal of Health Management 23*(3), 441–469.

WB (2012). *PKH conditional cash transfer: Social assistance program and public expenditure review 6*. World Bank.

WB (2017). *Towards a comprehensive, integrated, and effective social assistance system in Indonesia*. World Bank.

York, B., S. Loeb, and C. Doss (2018). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. *Journal of Human Resources 54*(3), 0517–8756R.

Zurovac, D., B. Larson, R. Sudoi, and R. Snow (2012). Costs and cost-effectiveness of a mobile phone text-message reminder programmes to improve health workers' adherence to malaria guidelines in kenya. *PLOS ONE 7*, 1–6.

8. Tables

Table 1: Overview of the Intervention (%)

Feature	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Entire period
	(1)	(2)	(3)	(4)	(5)
T: Anemia	0.38	0.23	0.31	0.23	0.29
T: Breastfeeding	0.08	0.31	0.15	0.15	0.17
T: Postnatal care	0.15	0.23	0.15	0.23	0.19
T: Vaccinations	0.31	0.15	0.15	0.23	0.21
T: Hygiene	0.08	0.08	0.23	0.15	0.13
Advise: No	0.00	0.03	0.00	0.03	0.01
Advise: External	0.69	0.33	0.44	0.46	0.48
Advise: Internal	0.31	0.64	0.56	0.51	0.51
Focus: Child	0.64	0.69	0.54	0.56	0.61
Focus: Mother	0.36	0.31	0.46	0.44	0.39
Frame: Neutral	0.31	0.38	0.49	0.49	0.42
Frame: Win	0.23	0.21	0.26	0.15	0.21
Frame: Loss	0.46	0.41	0.26	0.36	0.37
Content: New	0.90	0.67	0.38	0.21	0.54
Content: Reminder	0.10	0.33	0.62	0.79	0.46
# SMS	39	39	39	39	156

Notes: Sms related to anemia and hygiene practices comprised content that related to (i) the specific case of pregnancy/early childhood and (ii) general circumstances (independent of pregnancy/motherhood). 'External' advise refers to suggestions that involve being in contact with formal health service providers, while 'Internal' advise refers to suggestions that would change a mother's direct environment (e.g. home, shopping).

Table 2: Summary statistics: Outcome Variables at Baseline

Variable	Mean	Median	SD	Min	Max	Obs.
Panel A: All respondents						
Postnatal Care Practice	0.27	0.00	0.39	0.00	1.00	1821
Vaccination Practices (selfreported)	0.75	1.00	0.44	0.00	1.00	1821
Hygiene Practice	0.27	0.36	0.22	0.00	0.81	1821
Practice Index	0.43	0.46	0.23	0.00	0.94	1821
Mother has Anemia	0.32	0.00	0.47	0.00	1.00	1818
Child stunted	0.38	0.00	0.48	0.00	1.00	1714
Child wasted	0.12	0.00	0.33	0.00	1.00	1713
Anemia Knowledge	0.16	0.14	0.17	0.00	0.58	1821
Breastfeeding Knowledge	0.67	0.75	0.27	0.00	1.00	1821
Postnatal Care Knowledge	0.24	0.33	0.21	0.00	0.81	1821
Vaccination Knowledge	0.33	0.33	0.18	0.00	0.88	1821
Knowledge Index	0.35	0.35	0.13	0.00	0.71	1821
Panel B: Re-interviewed respondents						
Postnatal Care Practice	0.27	0.00	0.39	0.00	1.00	1726
Vaccination Practices (selfreported)	0.75	1.00	0.43	0.00	1.00	1726
Hygiene Practice	0.26	0.36	0.22	0.00	0.81	1726
Practice Index	0.43	0.46	0.23	0.01	0.94	1726
Mother has Anemia	0.33	0.00	0.47	0.00	1.00	1723
Child stunted	0.38	0.00	0.48	0.00	1.00	1625
Child wasted	0.13	0.00	0.33	0.00	1.00	1624
Anemia Knowledge	0.16	0.14	0.17	0.00	0.58	1726
Breastfeeding Knowledge	0.67	0.75	0.27	0.00	1.00	1726
Postnatal Care Knowledge	0.24	0.33	0.21	0.00	0.81	1726
Vaccination Knowledge	0.33	0.33	0.18	0.00	0.88	1726
Knowledge Index	0.35	0.35	0.13	0.00	0.71	1726

Notes: Postnatal Care Practice - Share of child health record books (presented to enumerator). *Vaccination Practice* - Child received at least one vaccination (self-reported). *Hygiene Practice* - Proportion of correct handwashing practices performed by respondent (observed by enumerator). *Practice Index* - average over all practice indicators related to postnatal care, vaccination and hygiene. *Mother has Anemia* - mother has moderate or severe anemia according to WHO (hemoglobin less than 9.0 g/dL). *Child Stunted* - youngest child is stunted (based on WHO standards). *Child Wasted* - youngest child is wasted (based on WHO standards). *Anemia Knowledge* - share of correct responses to eight knowledge questions related to anemia. *Breast Feeding Knowledge* - share of correct responses to four knowledge questions related to breast feeding. *Postnatal Care Knowledge* - share of correct responses to three knowledge questions related to post natal care. *Vaccination Knowledge* - share of correct responses to two knowledge questions related to vaccination. *Knowledge Index* - average over all knowledge questions.

Table 3: Impact on Behavior

	Postnatal Care Practices		Vaccination Practices		Hygiene Practices		Practice Index	
Panel A: Intention to Treat								
Assigned Treatment	0.077	0.075	0.068	0.060	0.105	0.096	0.084	0.076
	(0.035)**	(0.019)***	(0.028)**	(0.023)**	(0.021)***	(0.021)***	(0.018)***	(0.012)***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.028]	[0.000]	[0.000]	[0.000]
Panel B: Local Average Treatment								
Treatment Exposure	0.099	0.096	0.088	0.077	0.134	0.123	0.107	0.097
	(0.045)**	(0.024)***	(0.037)**	(0.030)**	(0.027)***	(0.027)***	(0.024)***	(0.016)***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.031]	[0.000]	[0.000]	[0.000]
Lagged Outcome	No	Yes	No	Yes(*)	No	Yes	No	Yes
Control Variabless	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.51		0.76		0.20		0.49	
Observations	1,726		1,726		1,726		1,726	

Notes: Panel A: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. Panel B: Local average treatment estimates using two-stage least squares instrumenting treatment exposure (whether mother reported to have received PKH sms on at least one topic) with assigned treatment status. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. In square brackets wyoung p-values controlling for the family-wise error rates implemented by `wyoung` (Jones et al., 2019). *Postnatal Care Practice* - Share of child health record books (presented to enumerator). *Vaccination Practice* - Child received at least one vaccination. *Hygiene Practice* - Proportion of correct handwashing practices performed by respondent (obersevd by enumerator). *Practice Index* - average over all practice indicators related to postnatal care, vaccination and hygiene. Controls include respondent characteristics (the respondent's baseline knowledge and practice index, her age, education, well-being, whether she is first-time mother, whether she is currently pregnant, her health related knowledge and practice at baseline, measured by the respective indices, and the number of children living in the household, the mother's degree of openness using the Big-5 personality measures, her bargaining power with respect to household and child related decisions, whether she regularly attends the monthly PKH meetings and her social network among the PKH mothers), village characteristics (the availability of a midwife in the village, whether the density of early childhood facilities as well as health centers exceeds the median in the sample, and whether the district capital is more than 1 hour traveling time away) as well as sub-district fixed effects. (*) The lagged outcome for vaccination practices is the self-reported vaccination practice, as it was verified only at endline.

Table 4: Impact on Health Outcomes

	Mother has Anemia		Child Stunted		Child Wasted	
Panel A: Intention to Treat						
Assigned Treatment	-0.096	-0.048	-0.013	-0.027	0.014	0.006
	(0.048)**	(0.023)**	(0.033)	(0.033)	(0.012)	(0.014)
	[0.046]	[0.037]	[0.694]	[0.415]	[0.231]	[0.666]
Panel B: Local Average Treatment						
Treatment Exposure	-0.123	-0.062	-0.017	-0.034	0.018	0.007
	(0.061)**	(0.029)**	(0.043)	(0.042)	(0.015)	(0.017)
	[0.047]	[0.036]	[0.694]	[0.415]	[0.231]	[0.666]
Lagged Outcome	No	Yes	No	n.a.	No	n.a.
Control Variabless	No	Yes	No	Yes	No	Yes
Control Mean	0.48		0.36		0.04	
Observations	1,720		1,492		1,492	

Notes: Panel A: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. Panel B: Local average treatment estimates using two-stage least squares instrumenting treatment exposure (whether mother reported to have received PKH sms on at least one topic) with assigned treatment status. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. In square brackets wyoung p-values controlling for the family-wise error rates implemented by `wyoung`. *Mother has Anemia* - mother has moderate or severe anemia according to WHO (hemoglobin less than 9.0 g/dL). *Child Stunted* - youngest child is stunted (based on WHO standards). *Child Wasted* - youngest child is wasted (based on WHO standards). Controls included as specified in Table 3. Lagged outcomes for stunting and wasting not available for children born after the baseline survey.

Table 5: Impact on Knowledge

	Anemia Knowledge		Breastfeeding Knowledge		Postnatal-care Knowledge		Vaccination Knowledge		Hygiene Knowledge		Knowledge Index	
Panel A: Intention to Treat												
Assigned Treatment	0.063	0.068	0.041	0.045	0.073	0.067	0.009	0.007	0.026	0.021	0.042	0.043
	(0.018)***	(0.017)***	(0.035)	(0.031)	(0.017)***	(0.016)***	(0.012)	(0.006)	(0.008)***	(0.008)**	(0.015)***	(0.012)***
	[0.001]	[0.000]	[0.245]	[0.154]	[0.000]	[0.000]	[0.483]	[0.254]	[0.003]	[0.012]	[0.004]	[0.001]
Panel B: Local Average Treatment												
Treatment Exposure	0.080	0.087	0.052	0.058	0.093	0.086	0.011	0.009	0.033	0.027	0.054	0.055
	(0.023)***	(0.021)***	(0.044)	(0.040)	(0.022)***	(0.021)***	(0.016)	(0.008)	(0.011)***	(0.011)**	(0.019)***	(0.016)***
	[0.001]	[0.000]	[0.244]	[0.154]	[0.000]	[0.000]	[0.483]	[0.254]	[0.003]	[0.013]	[0.005]	[0.001]
Lagged Outcome	No	Yes	No	Yes	No	Yes	No	Yes	No	n.a.	No	Yes
Control Variables	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.16		0.42		0.20		0.23		0.35		0.27	
Observations	1,726		1,726		1,726		1,726		1,726		1,726	

Notes: Panel A: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. Panel B: Local average treatment estimates using two-stage least squares instrumenting treatment exposure (whether mother reported to have received PKH sms on at least one topic) with assigned treatment status. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. In square brackets wyoung p-values controlling for the family-wise error rates implemented by `wyoung` (Jones et al., 2019). *Anemia Knowledge* - share of correct responses to eight knowledge questions related to anemia. *Breast Feeding Knowledge* - share of correct responses to four knowledge questions related to breast feeding. *Postnatal Care Knowledge* - share of correct responses to three knowledge questions related to post natal care. *Vaccination Knowledge* - share of correct responses to two knowledge questions related to vaccination. *Knowledge Index* - average over all knowledge questions. Controls included as specified in Table 3. Lagged outcome not available for hygiene knowledge as not elicited at baseline.

Table 6: Theory of Planned Behavior

	Postnatal Care Practices	Vaccination Practices	Hygiene Practices	Mother has Anemia	Child Stunted	Child Wasted
Panel A: Attitude – first child						
Treatment	0.067	0.042	0.097	-0.033	-0.014	0.003
	(0.019)***	(0.024)*	(0.022)***	(0.024)	(0.033)	(0.015)
First child	-0.017	-0.102	0.023	0.143	-0.018	-0.007
	(0.040)	(0.048)**	(0.032)	(0.048)***	(0.071)	(0.025)
Treatment x First child	0.059	0.130	-0.007	-0.118	-0.103	0.023
	(0.050)	(0.060)**	(0.038)	(0.057)**	(0.078)	(0.035)
Panel B: Attitude – openness						
Treatment	0.031	0.018	0.083	-0.053	-0.060	-0.006
	(0.023)	(0.028)	(0.025)***	(0.027)*	(0.040)	(0.020)
Open personality	-0.028	-0.016	0.009	0.003	-0.091	-0.013
	(0.020)	(0.025)	(0.014)	(0.031)	(0.038)**	(0.015)
Treatment x Open personality	0.092	0.089	0.029	0.010	0.065	0.025
	(0.031)***	(0.034)**	(0.028)	(0.040)	(0.048)	(0.022)
Panel C: Ability – negotiation power						
Treatment	0.066	0.056	0.102	-0.064	0.038	0.018
	(0.022)***	(0.027)**	(0.023)***	(0.028)**	(0.037)	(0.015)
High bargaining power	-0.013	-0.005	0.012	0.018	0.111	0.018
	(0.026)	(0.029)	(0.018)	(0.039)	(0.037)***	(0.016)
Treatment x High bargaining power	0.025	0.012	-0.016	0.040	-0.176	-0.031
	(0.034)	(0.038)	(0.029)	(0.048)	(0.055)***	(0.023)
Panel D: Norms – local practice						
Treatment	0.066	0.063	0.094	-0.003	-0.025	-0.000
	(0.033)*	(0.037)*	(0.023)***	(0.033)	(0.047)	(0.018)
Good local practice	-0.020	0.012	-0.000	0.020	-0.049	-0.014
	(0.038)	(0.038)	(0.025)	(0.035)	(0.046)	(0.020)
Treatment x Good local practice	0.010	0.017	0.002	-0.128	-0.058	0.006
	(0.043)	(0.058)	(0.036)	(0.057)**	(0.068)	(0.025)
Lagged Outcome	Yes	Yes(*)	Yes	Yes	No	No
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. The outcome variables are defined as in Tables 3 and 4. *First child* - is a dummy indicating that the respondent is a first-time mother. *Open personality* - is a dummy variable indicating that the mother has an above median score in openness defined based on the Big 5 taxonomy. *High bargaining power* - is a dummy variable indicating that the mother has an above median score in a bargaining index measuring bargaining power with respect to household and with respect to child related decision making. *Good local practice* - is a dummy variable indicating that the majority of respondents in a village have a practice index at baseline that is above the sample median. Controls included as specified in Table 3. (*) The lagged outcome for vaccination practices is the self-reported vaccination practice, as it was verified only at endline.

Table 7: Health Supply

	Postnatal Care Practices	Vaccination Practices	Hygiene Practices	Mother has Anemia	Child Stunted	Child Wasted
Panel A: Density of Health Centers						
Treatment	0.052	0.062	0.123	-0.034	-0.022	0.002
	(0.029)*	(0.036)*	(0.030)***	(0.025)	(0.043)	(0.016)
High density of health centers	0.009	0.069	0.065	0.033	0.020	0.006
	(0.034)	(0.040)*	(0.037)*	(0.043)	(0.056)	(0.026)
Treatment x High density of health centers	0.048	-0.004	-0.056	-0.030	-0.010	0.009
	(0.042)	(0.052)	(0.040)	(0.050)	(0.065)	(0.025)
Panel B: Density of Early Childhood Facilities						
Treatment	0.082	0.048	0.121	-0.093	-0.054	0.009
	(0.032)**	(0.040)	(0.029)***	(0.034)***	(0.045)	(0.017)
High density of early childhood facilities	0.001	0.045	0.004	-0.107	-0.070	0.002
	(0.038)	(0.040)	(0.024)	(0.042)**	(0.050)	(0.020)
Treatment x High density of early childhood facilities	-0.012	0.023	-0.045	0.082	0.049	-0.005
	(0.043)	(0.051)	(0.040)	(0.052)	(0.060)	(0.026)
Panel C: Midwife present in village						
Treatment	0.001	0.100	0.084	-0.016	-0.058	0.007
	(0.037)	(0.043)**	(0.023)***	(0.028)	(0.050)	(0.019)
Village midwife present in village	-0.060	-0.034	-0.010	-0.008	-0.006	-0.020
	(0.050)	(0.071)	(0.032)	(0.042)	(0.070)	(0.032)
Treatment x Village midwife present in village	0.114	-0.061	0.018	-0.050	0.047	-0.002
	(0.043)***	(0.056)	(0.037)	(0.042)	(0.062)	(0.026)
Lagged Outcome	Yes	Yes(*)	Yes	Yes	No	No
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. The outcome variables are defined as in Tables 3 and 4. *High density of health centers* - is a dummy variable equal to one if the number of health centers per 1,000 families is above the sample median. *High density of early childhood facilities* - is a dummy variable equal to one if the number of early childhood facilities per 1,000 families is above the sample median. *Village midwife present in village* - is a dummy variable equal to one if there is at least one trained village midwife present in the village. Controls included as specified in Table 3. (*) The lagged outcome for vaccination practices is the self-reported vaccination practice, as it was verified only at endline.

Table 8: PKH Structure

	Postnatal Care Practices	Vaccination Practices	Hygiene Practices	Mother has Anemia	Child Stunted	Child Wasted
Panel A: PKH involvement						
Treatment	0.071	0.085	0.093	-0.063	0.050	0.037
	(0.036)**	(0.041)**	(0.028)***	(0.039)	(0.056)	(0.023)
Regular PKH meeting attendance	0.020	0.048	-0.005	0.008	0.004	0.014
	(0.023)	(0.034)	(0.021)	(0.038)	(0.036)	(0.017)
Treatment x Regular PKH meeting attendance	0.005	-0.034	0.004	0.019	-0.103	-0.041
	(0.040)	(0.041)	(0.029)	(0.049)	(0.058)*	(0.026)
Panel B: Eligible PKH Mothers Network						
Treatment	0.051	0.065	0.084	-0.058	-0.018	0.013
	(0.021)**	(0.025)***	(0.022)***	(0.033)*	(0.039)	(0.016)
Large eligible Network	-0.010	0.000	-0.003	-0.003	0.067	-0.004
	(0.024)	(0.027)	(0.019)	(0.034)	(0.042)	(0.015)
Treatment x Large eligible Network	0.083	0.003	0.038	0.014	-0.031	-0.024
	(0.036)**	(0.044)	(0.030)	(0.050)	(0.059)	(0.020)
Lagged Outcome	Yes	Yes(*)	Yes	Yes	No	No
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Intention-to-treat estimates using OLS. Treatment indicating whether mother has been targeted by the sms campaign. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. The outcome variables are defined as in Tables 3 and 4. *Regular PKH meeting attendance* - is a dummy variable indicating that the mother joins the PKH meetings at least every two months. *Large eligible network* - is a dummy variable equal to one if the number of other PKH mothers eligible for treatment, who the respondent regularly visits (or is visited by) is above the sample median. Controls included as specified in Table 3. (*) The lagged outcome for vaccination practices is the self-reported vaccination practice, as it was verified only at endline.

Table 9: Overview on impact-cost estimates

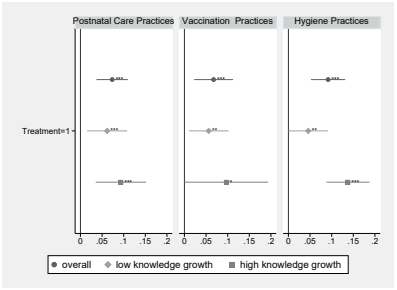
#	Program	Country	Indicator	Year of impact	Impact	Benefit	Cost 1	Cost 2	IC ratio
Panel A: Estimates on the impact-cost ratio of CCTs (own calculations based on literature review)									
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Bolsa Familia	Brazil	Child immunization	2010	2.3pp	520	761	1,029	0.002
	Bolsa Familia	Brazil	Post-natal visit	2010	3.1pp	520	761	1,029	0.003
2	Familias en Accion	Columbia	Child immunization	2003	6.0pp	186	377	472	0.013
	Familias en Accion	Columbia	Child stunting	2003	6.9pp	186	377	472	0.014
3	PKH	Indonesia	Child immunization	2009	3.9pp	180	380	586	0.007
	PKH	Indonesia	Ante-natal visits	2009	11pp	180	380	586	0.019
	PKH	Indonesia	Child stunting	2009	2.0pp	180	380	586	0.003
4	Academic CCT	Kenya	Child immunization	2018	3.1pp	92	313	313	0.010
	Academic CCT	Kenya	Child stunting	2018	7.0pp	92	313	313	0.022
5	Prospera	Mexico	Post-natal visits	2000	3pp	460	918	1,050	0.003
	Prospera	Mexico	Child stunting	2000	2pp	460	918	1,050	0
	Prospera	Mexico	Child anemia	2000	10.0pp	460	918	1,050	0.010
6	Academic CCT	Nigeria	Ante-natal visits	2018	14.2pp	14	42	42	0.333
Panel B: Estimates on the impact-cost ratio of mHealth interventions									
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	PKH SMS	Indonesia	Child immunization	2015	6.0pp	0	1.5	1.66	3.616
	PKH SMS	Indonesia	Post-natal care	2015	7.5pp	0	1.5	1.66	4.508
	PKH SMS	Indonesia	Hygiene	2015	7.6pp	0	1.5	1.66	4.568
	PKH SMS	Indonesia	Maternal anemia	2015	4.8pp	0	1.5	1.66	2.885
2	Academic RCT	Kenya	Correct malaria treatment	2010	24.9pp	0	7.8	13.59	1.84
3	Academic RCT	Ghana	Adherence to malaria treatment	2011	4.9pp	0	0.75	1.79	2.73
4	Academic RCT	China	Early breastfeeding	2011	4.8pp	0	1.56	1.80	2.672
5	Cohort study	Ethiopia	Reminder on antenatal visits	2014	3.6pp	0	0.60	0.88	4.078
6	Cohort study	Australia	Cholesterol levels	2012	4.5pp	0	9.6	10.77	0.418

Notes: 'Benefit' refers to the average benefit a beneficiary household receives in a year. It is measured in current USD at the time the impact was measured. 'Cost 1' refers to the total implementation costs in USD per beneficiary household (benefits + administrative costs) of the intervention at the time the impact was measured. 'Cost 2' adjusts the 'Cost 1' parameter to the common base year 2018 using national inflation rates. 'IC ratio' is the ratio between the stated impacts (column 6) and implementation costs per household (column 10). Please see section C.4 in the online appendix for more details on how estimates were derived.

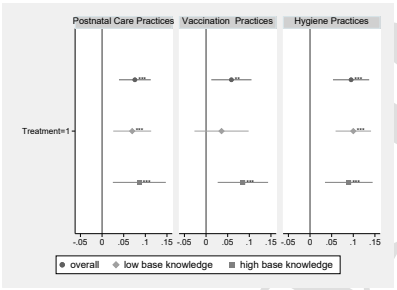
9. Figures

Figure 1: Treatment Effects by Knowledge Growth and Baseline Knowledge

Panel A: Impact on Behavior

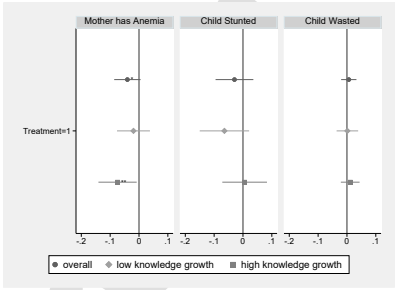


(a) by knowledge growth

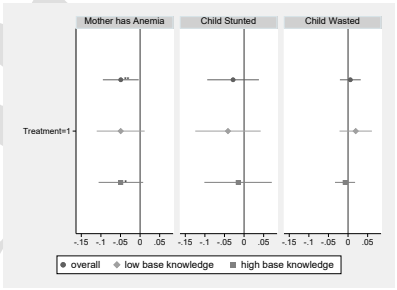


(b) by baseline knowledge

Panel B: Impact on Health Outcomes



(a) by knowledge growth



(b) by baseline knowledge

Notes: ITT estimates, controlling for mother and village characteristics and dependent variable at baseline (with the exception of stunting and wasting). Differential effects by (a) change in domain specific knowledge (above/below median) or (b) baseline knowledge (above/below median). As no knowledge information was elicited on hygiene practices and child health at baseline, overall knowledge index is used (i.e. average of postnatal care, vaccination and breastfeeding knowledge).

A. Appendix: Background Tables

A.1. Respondent and Village Characteristics

Table A.1: Summary statistics: Respondent and Village Characteristics (at baseline)

Variable	(1) Mean	(2) Median	(3) SD	(4) Min	(5) Max	(6) Obs.
Village Characteristics						
Rural area	0.43	0.00	0.50	0.00	1.00	127
Agriculture - village economic main sector	0.57	1.00	0.50	0.00	1.00	127
No of families living in village	2748.90	1003.00	3536.79	299.00	19099.00	127
Subdistrict capital - distance in hours	1.01	1.00	0.09	1.00	2.00	127
District capital - distance in hours	1.29	1.00	1.87	1.00	20.00	127
Strong phone signal	0.85	1.00	0.36	0.00	1.00	127
Posyandu - available in village	1.00	1.00	0.00	1.00	1.00	127
Health centers per 1,000 families	4.37	3.70	2.48	0.00	12.35	127
Early childhood facilities per 1,000 families	1.72	1.43	1.23	0.00	5.98	127
Village midwife present in village	0.67	1.00	0.47	0.00	1.00	127
No of PKH groups in village	2.85	2.00	2.05	1.00	13.00	127
No of PKH beneficiaries in village	38.94	31.00	30.73	5.00	184.00	127
Respondent and Household Characteristics						
Female	0.99	1.00	0.09	0.00	1.00	1821
Household head	0.02	0.00	0.15	0.00	1.00	1821
Married	0.98	1.00	0.14	0.00	1.00	1821
Muslim	0.92	1.00	0.26	0.00	1.00	1821
Age of person	31.98	31.00	6.58	15.00	72.00	1821
Years of education	7.16	6.00	3.38	0.00	15.00	1821
Household size	5.31	5.00	1.69	3.00	16.00	1821
No of children	2.98	3.00	1.60	0.00	14.00	1821
No of children in household $\leq$ 15 yrs	2.84	3.00	1.31	0.00	11.00	1821
First child	0.14	0.00	0.35	0.00	1.00	1821
Woman is pregnant	0.09	0.00	0.29	0.00	1.00	1821
Welfare index (PCA dwellings)	0.14	0.50	1.53	-3.36	1.87	1821
Subjective welfare	1.96	2.00	0.82	1.00	6.00	1821
Openness	0.03	-0.59	1.42	-4.52	3.20	1821
Bargaining power wrt household	0.91	0.94	0.10	0.00	1.00	1821
Bargaining power wrt child	0.87	0.91	0.12	0.00	1.00	1821
Phone Use	6.48	7.00	4.20	0.00	28.00	1821
PKH Network	8.15	8.00	1.97	2.00	13.00	1821
Attends PKH meetings on monthly basis	0.29	0.00	0.45	0.00	1.00	1821
Regular PKH meeting attendance	0.74	1.00	0.44	0.00	1.00	1821

Notes: Information is based on the sample selected for the RCT of 1,821 respondents living in 127 villages. For the description and coding of the variables see C.2 in Appendix C.2.

Table A.2: Balance of Respondent Characteristics at Baseline

Variable	(1) Mean Treatment	(2) Mean Control	(3) Treat. vs. Cont.
Aged 25 or younger	0.17	0.15	0.01
	(0.37)	(0.36)	(0.60)
Aged 26-35	0.55	0.55	-0.00
	(0.50)	(0.50)	(0.89)
Aged above 35	0.29	0.30	-0.01
	(0.45)	(0.46)	(0.79)
Years of education	7.04	7.27	-0.23
	(3.27)	(3.50)	(0.44)
First child	0.16	0.12	0.03
	(0.36)	(0.33)	(0.18)
No of children in household $\leq$ 15 yrs	2.77	2.90	-0.12
	(1.34)	(1.26)	(0.41)
Woman is pregnant	0.10	0.09	0.01
	(0.30)	(0.28)	(0.41)
Welfare index (PCA dwellings)	0.11	0.16	-0.06
	(1.56)	(1.52)	(0.81)
Subjective welfare	1.92	2.01	-0.09*
	(0.80)	(0.84)	(0.08)
Openness	0.02	-0.00	0.02
	(1.41)	(1.45)	(0.78)
Bargaining power (avg)	0.90	0.89	0.00
	(0.10)	(0.09)	(0.67)
PKH Network	8.26	8.06	0.20
	(1.91)	(2.04)	(0.56)
Regular PKH meeting attendance	0.70	0.77	-0.07
	(0.46)	(0.42)	(0.13)
High density of health centers	0.52	0.48	0.04
	(0.50)	(0.50)	(0.74)
High density of early childhood facilities	0.57	0.43	0.15
	(0.50)	(0.49)	(0.20)
Village midwife present in village	0.69	0.60	0.08
	(0.46)	(0.49)	(0.45)
District capital > 1hr away	0.02	0.06	-0.04
	(0.15)	(0.24)	(0.54)
Observations	883	843	1,726

Notes: Information is based on the sample of 1,726 respondents that were interviewed at baseline and endline. (1) and (2): standard deviations in parentheses; (3): p-values in parentheses with standard errors clustered on village level. */**/*** denote significance levels at 10/5/1 percent respectively.

Table A.3: Balance of Additional Village Characteristics at Baseline

Variable	(1) Mean Treatment	(2) Mean Control	(3) Treat. vs. Cont.
Rural area	0.41	0.45	-0.04
	(0.50)	(0.50)	(0.65)
Agriculture - village economic main sector	0.56	0.58	-0.02
	(0.50)	(0.50)	(0.80)
No of families living in village	2787.81	2710.59	77.22
	(3739.48)	(3354.54)	(0.90)
Subdistrict capital - distance in hours	1.02	1.00	0.02
	(0.13)	(0.00)	(0.32)
District capital - distance in hours	1.35	1.23	0.11
	(2.40)	(1.15)	(0.73)
Strong phone signal	0.90	0.80	0.11*
	(0.30)	(0.41)	(0.09)
Posyandu - available in village	1.00	1.00	0.00
	(0.00)	(0.00)	()
Health centers per 1,000 families	4.61	4.14	0.46
	(2.81)	(2.11)	(0.29)
Early childhood facilities per 1,000 families	1.71	1.73	-0.02
	(1.04)	(1.40)	(0.91)
Village midwife present in village	0.70	0.64	0.06
	(0.46)	(0.48)	(0.49)
No of PKH groups in village	2.95	2.75	0.20
	(1.96)	(2.15)	(0.58)
No of PKH beneficiaries in village	40.65	37.27	3.39
	(29.81)	(31.76)	(0.54)
Observations	63	64	127

Notes: (1) and (2): standard deviations in parentheses; (3): p-values in parentheses with standard errors clustered on village level. */**/*** denote significance levels at 10/5/1 percent respectively.

Table A.4: Balance of Outcome Variables at Baseline

Variable	(1) Mean Treatment	(2) Mean Control	(3) Treat. vs. Cont.
Postnatal Care Practice	0.25	0.29	-0.04
	(0.39)	(0.39)	(0.14)
Vaccination Practices (selfreported)	0.73	0.77	-0.04
	(0.45)	(0.42)	(0.19)
Hygiene Practice	0.27	0.26	0.02
	(0.22)	(0.22)	(0.19)
Practice Index	0.42	0.44	-0.02
	(0.24)	(0.23)	(0.24)
Mother has Anemia	0.32	0.33	-0.01
	(0.47)	(0.47)	(0.77)
Child stunted	0.38	0.38	0.00
	(0.49)	(0.48)	(0.99)
Child wasted	0.13	0.12	0.02
	(0.34)	(0.32)	(0.35)
Anemia Knowledge	0.16	0.16	0.00
	(0.16)	(0.17)	(0.94)
Breastfeeding Knowledge	0.66	0.67	-0.01
	(0.27)	(0.27)	(0.56)
Postnatal Care Knowledge	0.24	0.24	-0.00
	(0.21)	(0.21)	(0.89)
Vaccination Knowledge	0.34	0.33	0.00
	(0.17)	(0.18)	(0.77)
Knowledge Index	0.35	0.35	-0.00
	(0.12)	(0.13)	(0.84)
Observations	883	843	1,726

Notes: Information is based on the sample of 1,726 respondents that were interviewed at baseline and endline. (1) and (2): standard deviations in parentheses; (3): p-values in parentheses with standard errors clustered on village level. */**/*** denote significance levels at 10/5/1 percent respectively.

A.2. Message Handling

Table A.5: Summary Statistics

	mean	sd	N
Panel A: All Respondents			
Received sms on at least one topic			
Treatment	0.82	0.38	883
Control	0.06	0.24	843
Received sms on all topics			
Treatment	0.53	0.50	883
Control	0.00	0.07	843
Received sms and identifies PKH as sender			
Treatment	0.98	0.12	727
Control	0.48	0.50	52
Panel B: Treated Mothers who received a message			
Forwarded sms	0.04	0.19	741
Discussed sms	0.64	0.48	741
Discussed sms in PKH meeting	0.29	0.45	741
Read sms more than once	0.61	0.49	741
Saved sms	0.57	0.50	741
Happy about sms	0.85	0.36	741
Finds sms interesting	0.97	0.18	741
Finds sms important	0.98	0.14	741
Finds sms helpful	0.98	0.13	741

Notes: Based on self-reported data from the endline survey (N=1,726) .

A.3. Local Average Treatment Effect: First Stage

Table A.6: First Stage Results

	Treatment Exposure					
Analysis of:	*Behavioral Outcomes*		*Anemia Status*		*Stunting/Wasting*	
Treatment	0.781	0.781	0.781	0.780	0.785	0.784
	(0.016)***	(0.015)***	(0.016)***	(0.015)***	(0.016)***	(0.016)***
Observations	1726	1726	1720	1717	1492	1492
F Stat	2254.04	294.38	2266.31	301.31	2288.44	250.77
Control Variables	No	Yes	No	Yes	No	Yes

Notes: Outcome variable is *Treatment Exposure* (whether mother reported to have received PKH sms on at least one topic). OLS estimation for the full sample and different sub-samples used in the analysis. Treatment indicating whether a mother has been targeted by the sms campaign. In parentheses standard errors clustered on the village level. */**/*** denote significance levels at 10/5/1 percent respectively. Controls included as specified in Table 3.

Author Statement for the paper "Can health-information campaigns improve CCT outcomes? Experimental evidence from sms-nudges in Indonesia" (based on CRediT taxonomy)

Friederike Lenel: Methodology, Formal analysis, Writing – original draft, Writing – review & editing

Jan Priebe: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Project administration

Elan Satriawan: Conceptualization, Investigation, Resources, Writing – review & editing

Ekki Syamsulhakim: Conceptualization, Investigation, Resources, Writing – review & editing