The Impacts of a Multifaceted Pre-natal Intervention on Human Capital Accumulation in Early Life*

Pedro Carneiro Lucy Kraftman Giacomo Mason Lucie Moore
Imran Rasul Molly Scott†

November 2019

Abstract

We present results from a large-scale and long-term randomized control trial to evaluate an intervention targeting early life nutrition and well-being for households residing in extreme poverty in Northern Nigeria. The multifaceted intervention provides: (i) information to mothers and fathers on practices related to pregnancy and infant feeding; (ii) high-valued unconditional cash transfers to mothers, each month from pregnancy until the child turns two. We document two- and four-year impacts among 3600 pregnant women and their children. The intervention leads to large and sustained improvements in anthropometric and health outcomes for children, including an 8% reduction in stunting by endline. These impacts are partly driven by information-related channels (such as improved knowledge, practices and health behaviors of mothers towards new borns). However, the value and certain flow of cash transfers is also key: these induce labor supply responses among women, and allow them to undertake investments in livestock. These are both a source of protein rich diets for children, and generate higher earnings streams for households long after the cash transfers expire. The results show the sustainability and cost-effectiveness of scalable multifaceted pre-natal interventions in even the most challenging and food insecure economic environments.

JEL Classification I15, O15.

*We gratefully acknowledge financial support from DfID and the ESRC CPP at IFS (ES/H021221/1). Carneiro acknowledges the support of the ESRC for CEMMAP (ES/P008909/1) and the ERC (ERC-2015-CoG-682349). We thank Save the Children International and Action Against Hunger International, the OPM Abuja Survey team lead by Femi Adegoke, with special thanks to Ekundayo Arogundade, Eunice Atajiri, Joshua Moriyonu, Okechukwu Ezike, Tayo Ajala, Cynthia Nwodo, Ugochukwu Onuingbo, and Gloria Olisenekwu. Harold Alderman, Oriana Bandiera, Jere Behrman, Sonia Bhalotra, Adeline Delavande, Jessica Goldberg, John Hoddinott, Seema Jayachandran, Andrew Kardan, Michael Kremer, Ijaz Nabi, Kay Sharp, Alessandro Tarozzi, Marcos Vera Hernandez, Aly Visram and numerous seminar participants provided valuable comments. Human subjects approval was obtained on 2014-06-30 from the IRB at the National Health Research Ethics Committee of Nigeria (NHREC/01/01/2007-30/06/2014c). The study is registered on the AEA registry (AEARCTR-0000454). All errors remain our own.

†Carneiro: UCL, IFS and CEMMAP, p.carneiro@ucl.ac.uk; Kraftman: IFS, lucy.kraftman@ifs.org.uk; Mason: Competition and Markets Authority, giacomo.mason@cma.gov.uk; Moore: OPM, lucietafara@gmail.com; Rasul: UCL and IFS, i.rasul@ucl.ac.uk; Scott: OPM, Molly.Scott@opml.co.uk.

1 Introduction

Deprivation in early life has grave consequences for well-being through the life cycle [Almond and Currie 2011]. These consequences are particularly severe in terms of human capital accumulation, since physical and brain development are hindered by poor early life conditions [Doyle 2019]. Children growing up in extreme poverty are estimated to lose 25% of their income generating potential as adults [Richter *et al.* 2017]. Interventions to boost human capital accumulation in early life thus lie at the top of the policy agenda in poor countries. While such interventions have been shown to generate private, social and intergenerational returns, this evidence is largely derived from high- and middle-income settings. Yet the lowest levels of human capital accumulation are among children in Sub-Saharan Africa: two thirds of children in this region aged under five are stunted or growing up in extreme poverty [Grantham-McGregor *et al.* 2007]. There is thus an urgent need to understand whether welfare enhancing early life interventions can be implemented at-scale and cost effectively in this region.[1]

This paper presents evidence from a large-scale and long-term randomized control trial to evaluate an intervention designed to improve nutrition and well-being in the first 1000 days of life (including time *in utero*), by providing nutrition information and financial resources to a population with high rates of child malnourishment and extreme poverty. We show below that this program led to substantial improvements in the health and nutrition of targeted children, as well as remarkable impacts on the economic lives of mothers. These impacts can be observed well beyond the duration of the households' participation in the program.

The intervention, known as the Child Development Grant Program (CDGP), is implemented in Northern Nigeria. As Panel A of Figure A1 shows, Nigeria is the country with the highest absolute number of individuals living in extreme poverty (less than USD$1.90 per person/day). Infant mortality rates are 70 per 1000 births (Panel B), and the majority of children aged under five are stunted (Panel C). Our study context is an area of intense concentration of economic destitution within Nigeria: in our baseline sample, 85% of households live in extreme poverty, infant mortality rates are 90 per 1000 children, and 69% of children are stunted. Our sample villages are subject to frequent aggregate shocks, the agricultural cycle includes a lean season in which food is scarce and households have to resort to extreme coping strategies, and there are low levels of knowledge among women and their husbands about child-related practices.

The CDGP is a multifaceted intervention comprising a bundle of: (i) information provided to mothers and fathers on recommended practices related to pregnancy and infant feeding; (ii) high-valued unconditional cash transfers provided to mothers (beginning during pregnancy). The intervention thus simultaneously relaxes information and resource constraints. Panel D of Figure

[1] Impacts of various interventions in early life have been found on cognitive development and health [Campbell *et al.* 2014, Conti *et al.* 2016, Attanasio *et al.* 2019, Doyle 2019], schooling and labor market productivity [Hoddinott *et al.* 2008, Gertler *et al.* 2014] and across generations [Heckman and Karapakula 2019].

A1 illustrates how both are likely binding constraints in this setting. Using our baseline data, it shows the proportion of children, by household food expenditure decile, whose diet is comprised of one, two to three, or four or more food groups. Consuming four or more food groups is considered having a diverse diet. Although there is a gradient in dietary diversity by food expenditure decile, this gradient is small: 10% of households in the bottom decile have diverse diets, yet 5% of households in the highest decile have young children consuming just one food group. This suggests that a poor diet is not exclusively a result of lack of financial resources.

The CDGP is implemented at a village level, and is designed to be scalable: it trains locally hired community volunteers to deliver information messages and run the program day-to-day. The intervention is targeted to pregnant women, with the information provided covering pre-, peri- and post-natal stages of pregnancy. The value of the unconditional cash transfer is US$22 per month. This is substantial, corresponding to 85% of women's monthly earnings or 26% of monthly food expenditures. Women can start to receive transfers while the child is *in utero* until the child turns 24-months old (transfers are only provided for this child, not later borns).

The fact that women know transfers will be provided monthly in the first two years of the child's life provides them with a more stable flow of resources than is available from labor activities in these rural economies: the transfers almost act as a *de facto* temporary basic income for pregnant mothers. This opens up the possibility that they are used for both investment and consumption purposes. This is key because whether the cash component has short-lived or long-lasting impacts depends on this balance between investment and consumption.

We evaluate the intervention using a core sample of 3600 women pregnant at baseline. Two thirds of the 210 villages in our evaluation sample are randomly assigned to treatment. By focusing on women pregnant at baseline, we avoid issues of endogenous selection into pregnancy. We survey women, their husbands and gather information on mother-child interactions with a baseline survey, a two-year midline (that covers the critical window of the first 1000 days of life from conception), and a four-year endline. The timescale of our evaluation: (i) starts from before information and cash transfers are received, while the child is *in utero*; (ii) allows us to examine dynamic patterns of impact on children's health and human capital accumulation at two and four years post intervention; (iii) extends well after the cohort of women pregnant at baseline are actually in receipt of transfers, allowing us to understand whether the intervention has impacts on later born children, and whether the resource injection becomes self-sustaining if it is used to make investments that raise household income after two years.[2]

Our main results are as follows.

For the 'new' child (the one who was *in utero* at baseline) we start by focusing on height and stunting. Stunting is the best measure of the cumulative effects of chronic nutritional deprivation, reflecting the inability to reach linear skeletal growth potential, and is therefore a key indicator of

[2] We thus add to a nascent literature studying long run impacts of cash transfers in low-income settings [Baird *et al.* 2019, Bougen *et al.* 2019].

long-term well-being. We find large and persistent improvements in height-for-age Z scores (HAZ) for treated children over those in the control group. At midline, treated children have a statistically significant increase in their HAZ score by $.20\sigma$; (ii) at the lower tail of the distribution, there is a reduction in stunting of 8% (being below 2σ of the international norm); (iii) at the extreme lower tail of the distribution, there is a reduced incidence of extreme stunting of 14% (being below 3σ of the international norm). Most importantly, these impacts are sustained at endline, four years post-intervention, and well after cash transfers have stopped being disbursed. To get a sense of the magnitude of impacts, the midline ITT corresponds to the new child being .49cm taller, and the endline impact corresponds to being .62cm taller. As a benchmark, the mean difference in height between the top and bottom wealth quartiles is .24cm for children at two years of age.

For the CDGP intervention, we also see marked improvements in child health: there is a 12% reduction in illness/injury for children at midline, that improves slightly to 17% by endline. The incidence of diarrhea among the new child also falls dramatically: at midline there is a 18% reduction in the proportion of children experiencing an episode of diarrhea in the two weeks preceding the survey, with a 25% reduction by endline.

Given the multifaceted nature of the CDGP, our second set of results explore a sequence of potential mechanisms driving these child outcomes. These mechanisms can be divided between those predominantly related to the information components of the intervention, those predominantly related to the cash component, and those reflecting a combination of both.

On information-related mechanisms we find significant increases in knowledge among mothers and fathers. On each dimension of knowledge: (i) impacts on husbands are smaller in magnitude than for their wives; (ii) knowledge impacts are sustained at four years post-intervention, but relative to the control group, they do fade slightly over time. Of course, improvements in knowledge only translate to improvements in child outcomes if they are acted upon. We thus examine the peri-, ante- and post-natal practices women engage with in relation to their new child. Across all three dimensions of engagement, we find significant improvements in practices towards the new child in their first 1000 days of life: relative to Controls, mothers are 86% more likely to obtain antenatal care while the child was *in utero*, 37% more likely to put the new child to breast immediately, and almost three times as likely to exclusively breast-feed the new child for the first six months (as opposed to give them water, in a context where 27% of households use an unprotected dug well as their main water source).

Mothers' health behaviors towards the new child also improve dramatically at midline and endline. For example, the likelihood the new child is given deworming medication increases by 49% at midline, and by 74% at endline; the likelihood a child has received *all* their basic vaccinations increases threefold by endline. Even putting aside all the earlier documented impacts on child anthropometrics and health, these increases in deworming and vaccination rates in early life are likely to translate to long run impacts on children's lifetime welfare [Baird *et al.* 2016].

Transitioning towards mechanisms that reflect both information and resource components, we find the dietary diversity of foods consumed *specifically by* the new child improve at midline, and these impacts are sustained at four years. Moreover, we see large improvements in food security reported by households on survey date, greater food security reported across seasons (including the lean season, when food is scarce), and a reduced reliance on extreme forms of coping strategy to deal with food shortages.

The main sources of food driving increased dietary diversity are dairy products, flesh food and eggs, and other fruit/vegetables. The fact that two of these relate to produce derived from livestock is important for two reasons: (i) it links to other mechanisms related to the cash component of the CDGP, such as investments into business assets which we examine below; (ii) the consumption of such protein-rich foods early in life can drive physical growth and development in low-income settings [Dewey and Adu-Afarwuah 2008, Headey *et al.* 2018].

Focusing on mechanisms reflecting the transfer component of the intervention, we first examine labor market behavior. We find marked and permanent changes in the labor supply of women, in business investments of women (with no changes for men). The labor supply impacts for women occur on both the extensive and intensive margins, and reflect their increased engagement in self-employment activities in petty trading or livestock rearing. These changes in labor supply and investment in productive assets lead to earnings increases for women of 20% at midline, which are entirely sustained at endline, long after the CDGP transfers have expired. These changes are purely resource-related mechanisms: at no point of the intervention was it suggested to beneficiaries they should use transfers to engage in new forms of income generating labor activity or to undertake business investments.

Women's business inputs increase significantly by endline, long after cash transfers stop being disbursed to these women. We see no corresponding increase in expenditures on inputs for the husband's business. On livestock, women's ownership of any animal increases by 6.6pp (8%) at midline, rising significantly to 12.1pp (15%) at endline. Livestock ownership is critical in this economic environment because: (i) it generates earnings for women from the sale of animal produce such as milk and eggs; (ii) it produces a stable earnings stream all year round, thus reducing the volatility of womens' earnings; (iii) animal produce can be consumed at home, and this maps closely to the documented impacts on dietary diversity of the new child.

Monthly food consumption rises by $21 (25%) at midline, and this increase is mostly sustained at endline. By endline, the stock of household savings increases, and the stock of outstanding borrowings fall.

Pulling together these strands in a household budgeting exercise, we see an increase in net resources available to the household. The magnitude of the increase is $49 at midline, so more than double the value of the cash transfer itself ($22). In other words, the program induces large behavioral responses of household members, that not only improve the anthropometric and health outcomes of the new child, but that endogenously generate higher resource flows into the household overall. This increase in net resources is sustained at endline because the loss of transfers from

CDGP is offset by an increase in earnings and net savings. As a result, by endline we find a 5% reduction in extreme poverty rates among beneficiary households. This is a dramatic reduction achieved over relatively short period, by an intervention predominantly designed to improve early life nutrition.

We use mediation analysis to shed light on the relative roles of information- and resource-related channels in mediating anthropometric and health outcomes for the new child. Our analysis provides insights on the dynamics of behavioral change induced by the intervention, by examining how mechanisms vary at midline and endline. The results suggest that for HAZ outcomes, information-based mediators are more important in the new child's first 1000 days of life, explaining around 60% of the total ITT impact at midline, but the relative importance of resource-based channels increases dramatically between midline and endline.

By documenting the delicate and dynamic interplay between information- and resource-based mechanisms, we add to and considerably extend a nascent literature explicitly examining multi-faceted interventions to drive human capital accumulation in early life [Levere *et al.* 2016, Fernald *et al.* 2017, Ahmed *et al.* 2019].[3]

Our final set of results assess the cost effectiveness of the intervention. We do so in two steps. First, only accounting for the impact on the endogenous increase in net resources to the household (over and above the value of the cash transfers), the internal rate of return (IRR) to the program is over 200% even if net resource impacts die out after five years. Second, we then focus only on the monetary gains of increased height through earnings, exploiting estimates of the height-earnings gradient estimated in the longitudinal study of Hoddinott *et al.* [2013]. Doing so we estimate an IRR of 5.8% for boys and 3.8% for girls. Of course, this grossly underestimates the true return because we place no value on gains coming from non-earnings sources (and earnings gains only start to be accumulated once the child turns 16 while intervention costs are borne up front). However, under conservative assumptions of the short run (pre-labor market) gains to children of the intervention, the IRR to the program can rapidly increase to closer to 20%, comparable to

[3] This evidence is however not from Sub-Saharan Africa. Levere *et al.* [2016] use an RCT based in Nepal using county-level randomization to contrast the impacts of information, cash, and information plus cash on children in poor families. Relative to our study, they study lower-valued cash transfers ($7 per month) that last for less time (5 months). They combine impacts on pregnant mothers and those that already have a child when the program starts. Due to differential attrition at midline, they focus their study on endline impacts measured 18 months post-intervention. They find only the combined intervention impacts child cognition, but find no impact on anthropometrics. The show mechanisms related to maternal knowledge and practices, but not on channels related to labor supply, investment or earnings. Ahmed *et al.* [2019] present evidence from a similarly designed two-year experiment in Bangladesh that provided households with high valued cash transfers ($19 per month) until the child turned two, in-kind transfers of food, a combination of the two, cash plus information, and food plus information. The information on child-related practices was similar in design to the CDGP. They find *only* the combination of cash plus information significantly impacted HAZ and reduced stunting by 7.8pp. The paper highlights increased dietary diversity as a key channel, especially the consumption of protein-rich animal produce. As in Levere *et al.* [2016], it does not document channels related to labor supply, investment or earnings. Fernald *et al.* [2017] evaluate a combined group-based parenting classes and cash transfer program against just cash, in a sample of households in rural Mexico. They also find only the multifaceted intervention impacts child development, driven by impacts among indigenous households.

estimates of early life interventions in high-income settings where a fuller range of benefits can be monetized [Heckman *et al.* 2010].

Our contribution is to provide a large-scale and long-term evaluation of a scalable intervention to foster human capital accumulation among the poorest households in a context in Sub-Saharan Africa. In doing so we help build the evidence base for pre-natal interventions in exactly the context where early life deficits are most acute.

In the earlier literature, cash transfers have been argued to impact child outcomes through two channels. The first is a standard income effect: households will use additional resources to purchase inputs to raise child development if it is a normal good. The second channel highlights that poverty increases stress, and reduces the quality of cognitive reasoning, that can lead to worse parenting practices [Mani *et al.* 2013, Dean *et al.* 2018]. What has been far less emphasized or documented are the labor supply, investment and earnings channels that can follow from large and sustained cash transfers that provide a *de facto* basic income for women with young children. By going well beyond a focus on the first 1000 days of life, we show that the stable flow of transfers provided is critical for driving sustained impacts on children's human capital: they lead to transformative effects on household resources as women beneficiaries expand labor activities and make asset investments. Households are thus able to transform short run cash transfers into sustained endogenous changes in income. As a result, the effect size impacts on the human capital accumulation we find are at the upper end of documented impacts of early life interventions in middle-income contexts.[4]

Taken together out findings show the cost effectiveness and sustainability of scalable pre-natal interventions in even the most challenging and food insecure economic environments.

Section 2 details the intervention, data and design of the experiment. Section 3 presents ITT impacts on child outcomes for anthropometrics, health, cognitive and non-cognitive development. Section 4 examines information- and resource-related mechanisms driving these impacts, and conducts a mediation analysis to provide insights on the relative importance of these channels over time. Section 5 presents and internal rate of return calculation for the intervention. Section 6 concludes with the broader implications of our study for policies designed to foster human capital accumulation in early life in settings of extreme poverty, food insecurity and economic volatility. The Appendix presents further results on impacts on maternal health, within household spillovers to older and younger siblings.

[4]For example, in conditional cash transfer programs that have been prevalent in middle-income contexts, conditionality often requires households to undertake some positive parenting practices [Sridhar and Duffield 2006]. Maluccio and Flores [2004] present evidence from a conditional cash transfer program in Nicaragua and find two-year impacts on HAZ of $.13\sigma$. Macours *et al.* [2012] also present evidence from Nicaragua of a conditional cash transfer program and find two-year impacts on HAZ of $.07\sigma$, but these fade out after four years. The evidence from recent systematic reviews on the impact of cash transfers alone on child anthropometrics is even less encouraging [Manley *et al.* 2013, Caeyers *et al.* 2016], while systematic reviews of information-based interventions suggest the largest reductions in stunting occur in food secure populations [Bhutta *et al.* 2008, 2013].

2 Intervention, Data and Experimental Design

2.1 Program Design and Context

The Child Development Grant Programme (CDGP) is a multifaceted intervention comprising a bundle of: (i) information provided to mothers and fathers on recommended practices related to pregnancy and infant feeding; (ii) unconditional cash transfers to mothers.

Our evaluation is based on 210 villages in two states in North West Nigeria: Zamfara and Jigawa. Households are almost entirely of Hausa ethnicity and Muslim religion, and are structured around a male household head. Women are often secluded during daytime but engage in income-generating activities such as petty trading or rearing livestock, and they retain control over income streams they generate. As shown below, there is very limited knowledge of child nutrition practices, and the majority of households reside in extreme poverty and lack resources to fully invest in children's human capital. As a result, at baseline over two thirds of children under 5 in eligible households are stunted (one third are severely stunted).

The CDGP is provided at the community level and is targeted to pregnant women. The information provided thus covers pre-, peri- and post-natal stages of pregnancy, and women can start to receive transfers while the child is *in utero*. Given the role maternal nutrition and behavior during pregnancy plays in child growth and development, the intervention might have greater returns than programs starting post-natally [Currie and Almond 2011, Bhutta *et al.* 2013].[5]

The intervention is designed to be scalable within Nigeria and portable to contexts through Sub Saharan Africa with low state capacity: the day-to-day running of the program is the responsibility of locally-hired community volunteers (CVs). CVs can be of two types: (i) a lead CV (one per village), that is typically a skilled individual, that is further trained in a specialized counselling role; (ii) nutrition promoter CVs (two per village), who disseminate information on recommended practices and refer women to more senior CDGP staff when necessary. The lead CV is paid, while the nutrition promotion CVs receive a stipend to cover transport and meals, and certified training for their role. Administrative records show both types of CV work for around 25 hours/month.

Information Information messages are tailored to the context. They were developed by our intervention partners – Save the Children (SC) and Action Against Hunger (AAH) – to tackle prevalent and important knowledge gaps among the rural poor.[6] Panel A of Table A1 shows the eight key messages disseminated, covering practices of child care and nutrition during the pre-,

[5] In rural Nigeria, communities are normally subdivided into traditional wards, that represent a community subdivision made up of a separate cluster of households. In cases where communities were too large to serve as sampling units, we randomly selected one ward in the community. In cases where a sampled community had less than 200 households, we merged it with the neighboring community. We refer to these sampling units as villages.

[6] The CDGP program is implemented in Zamfara by SC, and in Jigawa by AAH. The exact same program is implemented by both NGOs, using common modalities. The evaluation takes place in five LGAs in these two states: Anka, Tsafe in Zamfara, and, Buji, Gagarawa and Kiri Kasama in Jigawa.

peri- and post-natal periods. Messages also encourage mothers to increase their food intake during pregnancy, and emphasize good hygiene and sanitation. These messages were developed based on an earlier nutritional intervention conducted by SC in Northern Nigeria, and gathering qualitative and quantitative information from stakeholders including households with young infants, community health workers, traditional birth attendants, and traditional/religious leaders, as well as guidelines issued by the Nigerian Federal Ministry of Health. Figure A2 provides an example of the visual aids used by CVs to convey messages.

Panel B of Table A1 details how information messages are delivered. Low-intensity channels include posters, radio, Friday preaching/Islamic school teachers, health talks, food demonstrations, and pre-recorded SMS/voice messages. The food and health demonstrations are delivered by trained CDGP staff, assisted by the CVs. They take place each month in each village. These low-intensity channels represent a 'one-size-fits-all' approach to communication, where individuals are passive recipients of messages. The intent is to provide information beyond those immediately eligible, including women likely to become pregnant in future, and to others influential in village life including men and older women. The latter group are especially important to target because they are the conventional source of information for pregnant mothers seeking advice on pregnancy and infant feeding [Sharp *et al.* 2018].

High intensity channels (offered in a random subset of villages as we discuss below) include small group parenting sessions (focusing on nutrition and health practices), and one-to-one counselling in home visits.

Cash Transfers The value of the unconditional cash transfer – US$22 per month (at the PPP exchange rate in August 2014) – was calibrated by our intervention partners to correspond to the cost of a diverse household diet (not accounting for any crowd out of existing food expenditures). However benchmarked, the value of the monthly transfer is substantial: at baseline, it corresponds to 12% of household monthly earnings, 85% of women's monthly earnings, or 26% of monthly food expenditures. Moreover, the fact that it is known that transfers will be provided each month until the child is 24 months old provides women with a more stable flow of resources than is available from most labor activities in these rural economies. The magnitude and certainty of transfers opens up the possibility that they are used for both investment and consumption outcomes.[7]

We might think of this as a labelled cash transfer given it is bundled with information on child-related practices, nutrition, health and sanitation [Benhassine *et al.* 2015]. However, at no point of the intervention was it suggested to beneficiaries that they should use the cash transfers to engage in new forms of income generating labor activity or to undertake business investments.

Women had to meet two criteria to be eligible: (i) be resident in a village in which the CDGP

[7]The value of the cash transfer increased from NGN3500 to NGN4000 from January 2017 onwards. This later change is not relevant for the core sample of women pregnant at baseline that we focus on. Throughout our analysis, all monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

was implemented; (ii) be pregnant, as verified by an on-the-spot urine test in the presence of a female CV [Sharp *et al.* 2018]. Once eligibility was established, thumbprints were taken to be used when transfers were disbursed.[8] Conditional on meeting these criteria, the program is universal, avoiding any costly verification of a household's poverty status. As soon as women were deemed eligible, they could begin receiving transfers. These were provided each month until the child was 24 months old.

Cash transfers were delivered by payment agents who visited villages monthly, using thumbprints to identify the correct eligible women, and transferring cash directly to them. Women are eligible to receive transfers for one child only – the child *in utero* when eligibility is established.[9]

2.2 Timeline and Data Collection

The intervention was piloted between April and July 2014 to iron out implementation difficulties, and then scaled-up for this evaluation. Figure 1 shows the timeline of activities from June 2014 in the 210 villages in the evaluation. Villages underwent a one week period of intense mobilization, involving local and religious leaders, where the CDGP was implemented. The low-intensity information channels of the CDGP serve as a continuation of the village mobilization.[10]

We conducted a village census covering 38,803 women aged 12-49 in the 210 villages. 83% of them were married, 53% were in polygamous relationships. This census allows us identify households with a pregnant woman, and so immediately eligible for the program, as well as women aged 12-49 at risk of becoming pregnant during the evaluation period.[11] Our baseline survey took place from August to October 2014, our midline survey was conducted in October/November 2016, and the endline survey took place from August to October 2018.[12]

Surveys and Sampling From the census we drew a sample of 26 women per village and their husbands. Each is interviewed separately on survey modules covering knowledge related to pregnancy and infant nutrition, infant and young child feeding practices, as well as consumption, savings/borrowing, asset ownership/investments, and their labor activities. This allows us to

[8] Once eligibility is confirmed, women are enrolled in an electronic database used for cash payments. Women are provided a mobile phone and a recharge card required to activate it. The mobile number acts as their unique ID in CDGP administrative records. It was originally planned for the phones to be used for mobile payments, but this proved infeasible. In practice, the phones are used primarily as to alert beneficiaries about payment dates.

[9] In the case of maternal mortality, payments would still be disbursed to a female caregiver of the child. In the case of child mortality, the women remain eligible for a later child. Finally, for polygamous households, multiple wives in the same household can be eligible.

[10] Given low levels of state capacity in North West Nigeria, there remained some variation in implementation quality: this was mainly driven by logistical supply-side issues (staffing, procurement), and caused delays in information provision in Jigawa. Cash transfers began to be disseminated from August 2014 onwards.

[11] Households are defined as individuals residing in the same dwelling unit with common cooking/eating arrangements. Polygamous husbands can rotate dwellings where they sleep, as wives are not always in the same dwelling.

[12] The lean season in rural North West Nigeria runs from March to October: this is when food is in short supply and households have sometimes to resort to extreme coping strategies. This coincides with the baseline and endline surveys, but this timing does not differ between treatment and control villages.

build a detailed picture of the information- and resource-based mechanisms linking the program components to child outcomes.

Our baseline sample covers 5433 women and 5413 husbands. This includes two types of women: (i) all those who are pregnant; (ii) those predicted most likely to become pregnant in the first two years of the evaluation.[13] Of the 5433 women surveyed at baseline, 3688 (68%) are pregnant at baseline. This evaluation focuses on this cohort of women, thereby avoiding issues of endogenous selection into pregnancy due to the program, and endogenous responses to the announcement of the program ending in the final year of our evaluation (as Figure 1 shows). The latter sample of women were surveyed to study such endogenous responses. Among the cohort of women pregnant at baseline, we refer to the child *in utero* at baseline as the 'new' child. The new child is the one for whom the cash transfer component of the CDGP is provided until she is 24 months old.

Over the course of the evaluation we implemented mother-child specific surveys to collect anthropometric, nutrition, health and developmental related outcomes for three different children.

First, at baseline we fielded a mother-child questionnaire to collect information about a randomly selected child aged 0-60 months. We refer to this as the 'old' child. We surveyed 2597 old children of women pregnant at baseline. Second, at midline we added a survey for the new child (the one *in utero* at baseline). Of the 3688 women pregnant at baseline: (i) 5% had no new children by midline; (ii) 83% had one new child; (iii) 12% had more than one new child. If a woman had more than one child since baseline, we randomly selected one of their children aged 0-2 at midline. We surveyed 2718 new children at midline. Finally, it is possible that between midline and endline, another child is born after the new child (their younger sibling). Hence at endline we implemented a mother-child questionnaire to also collect information on these children, whom we refer to as the 'end' child. We surveyed 2719 end children at endline.

We use our old and end child surveys to measure spillovers of the program onto older and younger siblings of the new child. Of course, the old child is present when information and cash transfers are being received. By the time the end child is born, their outcomes can be impacted through persistent gains in knowledge and practices generated by the information component, and any sustained increases in household income generated by the earlier receipt of cash transfers.

Finally, we administered community surveys to focus groups of village elders to record village-level shocks, infrastructure, the provision of health services from local clinics, and the provision of other programs.

[13] To predict which women were most likely to become pregnant between baseline and midline, we use the Nigeria DHS 2013 data to predict the likelihood of becoming pregnant using the covariates common with our household census: age, time since last birth, household size, number of children aged below/over five and TV ownership. The estimated coefficients from a LPM on the DHS data is then used to predict the pregnancy probability in the CDGP census data. The mass of predicted probabilities lies above .5. We selected non-pregnant women with the highest predicted probability, and we note the average predicted probability is the same in treatment and control villages.

2.3 Randomization and Treatments

Villages were randomly assigned to a control group or one of two treatment arms. These treatment arms varied only in the intensity of information delivered (the cash component of the program was identical). The first treatment arm (T1) provided information through the low-intensity channels described above: education posters, radio, Friday preaching/Islamic school teachers, health talks delivered by trained health workers, food demonstrations delivered by CDGP trained staff, and pre-recorded SMS/voice messages. The second treatment arm (T2) provided the same low-intensity information channels, but in addition offered: (i) infant and young child feeding support groups made up of 12-15 mothers, led by a CV and health extension worker, and meeting for an hour each month; (ii) one-to-one counselling for mothers with CVs – initiated at the request of women, or if the CV identified the need for a home visit. For cases involving issues beyond the scope of the CV (e.g. mastitis), the mother would be referred to a health facility. Overall, T2 provided more interactive and customized communication to mothers than T1.

We divided the villages into three tranches, with random assignment of villages taking place within each tranche or strata. This is because of the need to have the program implemented soon after pregnant women had been identified so that information flows and cash transfers could begin while their child was *in utero*. Given low levels of state capacity, there were some logistical delays in setting up transfer payments. As Figure 1 shows, transfers began being disseminated in August 2014, some three to four months after registration took place and information provision through the low-intensity channels began.

2.4 Attrition, Balance and Sample Characteristics

By the four-year endline, 23% of women pregnant at baseline had attrited. Table A2 shows that attrition is: (i) uncorrelated to treatment; (ii) almost perfectly predicted by whether the village is insecure (and thus enumerators were unable to travel there and interview *any* households) – indeed, in villages that were always secure, only 8% of pregnant women attrit by endline; (iii) there is no evidence of differential attrition in treated villages by baseline characteristics of women or their households (Column 3): the p-value on the joint significance of these interaction is .503. Columns 4 to 6 show similar levels and correlates of attrition for husbands of women pregnant at baseline, the old child (that is tracked between baseline and midline), and the new child (that is tracked from midline to endline).[14]

Table 1 shows balance by Control, T1 and T2 groups. Given the rolling enrolment and randomization tranches, the samples are well balanced on village and household characteristics, as well as characteristics of pregnant women and their husbands.

[14] At midline, enumerators were unable to visit 18 villages due to security risks, and this rose to 28 villages at endline. Village insecurity is itself not correlated to treatment, but largely relates to various types of man made shock that the village experiences such as curfews, violence, or widespread migration into the village.

This provides useful detail on the study context. Panel A shows village characteristics excluding villages our enumerators were unable to reach due to security risks. The remaining secure villages are still vulnerable to aggregate shocks: over 80% have been hit by a natural shock in the year prior to baseline (such as crop damage caused by weather or pests, floods and droughts), with at least a third having been hit by a man made shock (such as curfews, violence, or widespread migration into the village). With such a high degree of background uncertainty, the prospect of receiving substantial cash transfers each month for the first two years of the new child's life provides a great opportunity for households to invest such resources for longer term gains, as well as for immediate consumption. The features of our context and program almost make the cash transfers provided a form of temporary basic income for pregnant mothers.

Despite being poor rural locations, villages are in close proximity to markets and health facilities. Indeed, on average health facilities are 1.5km away, suggesting resource costs of travel are unlikely to explain their low usage rates.

Panel B shows that there are on average 7 individuals per household. Monthly food expenditures are $85 (whereas the monthly CDGP transfer is $22). Around 40% of monthly expenditures are on food. 70% of households live in extreme poverty, below the $1.90/day global threshold. They also suffer food insecurity, with 15% reporting not having enough food at some point during the year. The lean season in rural North West Nigeria runs from March to October: this is when food is in short supply and households have to sometimes resort to extreme coping strategies.

Panels C and D show characteristics of women pregnant and their husbands, at baseline. Despite women being age 25 on average, they have 4.6 children alive, aged below 18 and resident with them. Around half are in polygamous marriages with far older husbands (they are on average aged 43). Both spouses have low levels of human capital, with 20% of women being literate, and 40% of men being literate. The main labor activity for women is to rear/tend or sell household livestock: 36% are engaged in such work. Among men, over 80% have farming household land as their main labor activity.

Table 2 examines balance on child related outcomes. We again see that on most dimensions, samples are well balanced across Control and Treatment groups.

Parental knowledge on child nutrition practices is generally inadequate (especially with regard to early breast-feeding initiation, exclusive breast-feeding, and appropriate complementary feeding). For example, Panel A shows that only 14% of pregnant women believe a child should be exclusively breast-fed for the first six months of life (and thus are likely to provide the child water instead, in a context where 27% of households use an unprotected dug well as their main water source). Panel B shows husband's knowledge is equally low at baseline, so there is ample scope for both spouses to learn from the information provided in the CDGP.

Panel C shows characteristics of the old child (who on average are 36 months old at baseline). The low levels of knowledge are reflected in actual behaviors: only one third of these children were put to breast in the first 30 minutes after birth, and only a quarter were appropriately breast-fed (this is a dummy indicating age-appropriate breast-feeding according to WHO guidelines [WHO 2008], i.e. exclusive breast-feeding up to the age of 6 months and complementary breast-feeding from 6 to 23 months). On health outcomes, mothers report a third of children having diarrhoea in the two weeks prior to baseline. Finally, we find remarkably high levels of stunting: around two thirds of old children are stunted (so with a height-for-age Z-score (HAZ) below -2 standard deviations of the WHO defined guidelines [WHO 2009]). However, the incidence of wasting is lower in this population, so malnourishment shows up in height rather than weight.

Against this background of low levels of knowledge and poor practices, the information provided by the CDGP to mothers and fathers can have drastic impacts on child well-being.

Finally, Panel D relates to the new child – that is *in utero* at baseline. Based on mother's self-reports, these children are in the fifth month of pregnancy at baseline: hence the information and resource injections provided start from the last trimester of pregnancy onwards.

2.5 Take-up

We derive take-up rates for the cash transfer component of the CDGP using administrative records. Panel A of Table 3 shows that, in treatment villages, over 90% of households with women pregnant at baseline (and so immediately eligible for transfers) received payments by midline. The primary reason for not taking up is that women were initially misclassified as being pregnant (this applied to 42% of women that do not take-up by midline). We also note a small degree of take-up in Control villages (11%): this is likely due to cross-village registrations and implementation errors.

Panel B focuses on the timing of payments: on average, women start receiving cash transfers in their final month of pregnancy. 40% receive their first transfer sometime during pregnancy, 12% start receiving them in the month of birth, and 33% start receiving them post-natally.

Panel C measures treatment intensity: by midline, women have received on average 24 payments, of cumulative value $470. This corresponds to 2.5 months of household earnings. Around 84% of households are still receiving payments at midline for the new child that was *in utero* at baseline: the others are not receiving payments largely because of child mortality. However, the majority of women become pregnant again before endline, with 9% of them receiving payments by then (for their first surviving child, in line with eligibility conditions).

Overall we see no differences between T1 and T2 in the take-up of cash transfers, payment start dates or the duration of payments received.

On information, the low-intensity channels used in T1 and T2 provide information as a public good, so all households will be exposed to some extent. This is confirmed in Table A3 that shows 90% of women and husbands in treated villages report being exposed to at least one message via a low-intensity channel. There are message spillovers into Controls (as expected given that radio messaging is used), but reassuringly, only 4% of women in Control villages report receiving information from all low-intensity channels, while this rises to over 20% for women in Treated villages. Between genders, women are significantly more likely to be exposed to low-intensity channels than their husbands; for neither spouse are there significant differences between T1 and T2.[15] Panel B shows reports on exposure to the high-intensity channels: there remain large differences between Control and Treated villages, although the differences between T1 and T2 are not very marked. It is difficult to determine whether this is because low- and high-intensity channels are confused for each other, or whether the more intense forms of communication were poorly implemented.

Figure 2 shows descriptive evidence on the recall of each of the eight key messages provided, as measured at midline. The top panel shows this for women pregnant at baseline, and the bottom panel does so for their husbands. Table A4 shows the corresponding statistics and tests of equality across treatments and spouses. The data from the Control group shows there are real knowledge deficits among both spouses, and low levels of human capital among children are unlikely to only reflect resource constraints preventing households from implementing recommended child-related practices. We see that: (i) for all eight key messages, both treated spouses have significantly higher recall than individuals in the Control group; (ii) women have significantly higher rates of recall than husbands; (iii) for most messages, there are no significant differences in recall across spouses or between T1 and T2.

Taken together, the evidence shows little differential exposure to low- and high-intensity information channels. Hence we pool T1 and T2 treatment arms together for our analysis.

2.6 Empirical Method and Measures

We use the following ANCOVA specification when considering outcomes of mothers, fathers and the old child because these are all measured at baseline:

$$Y_{ivt} = \gamma_M T_v.(1 - E_t) + \gamma_E T_v.E_t + \alpha Y_{iv,t=0} + \beta X_{iv,t=0} + \eta_d + \lambda_s + \omega E_t + \varepsilon_{ivt}. \tag{1}$$

Y_{ivt} is the outcome, T_v is a treatment indicator, E_t is an endline wave indicator, $(1 - E_t)$ is a midline wave indicator, $X_{iv,t=0}$ are baseline controls, η_d is a district (LGA) fixed effect, λ_s are randomization strata (the tranches used given rolling enrolment into the program). ε_{ivt} is clustered by village given this is the level of the intervention. For outcomes related to the new child who was *in utero* at baseline (and for the end child that is only observed at endline), we obviously cannot control for $Y_{iv,t=0} = 0$ (the baseline outcome).[16]

(γ_M, γ_E) are the coefficients of interest: the two- and four-year intent-to-treat impacts of the CDGP intervention.

[15] The lower exposure of husbands to low-intensity channels is driven by them not attending food demonstrations.

[16] The baseline characteristics of the household and mother in $X_{iv,t=0}$ are the number of children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships, and the gender of the new child.

Where necessary, we present p-values adjusting for multiple hypothesis testing using stepdown adjustments [Romano and Wolf 2005]. In other cases, it is more appropriate to construct summary indices from a group of indicators using the method of Anderson [2008]. This uses the data covariance matrix to construct a weighted sum of indicators in the group, and so gives less weight to items more correlated with each other. These indices are standardized to have mean zero and variance one in the Control group.

The main set of outcomes we consider for the new child relate to their anthropometrics. To minimize measurement error, this information was collected by a dedicated anthropometric enumerator in each survey wave. We record child i's height, weight, and middle upper arm circumference. We use these to derive age-normed indicators of child development and nutritional status. We focus on height-for-age Z-scores (HAZ), as these relate to stunting: stunting is the best measure of cumulative effects of chronic nutritional deprivation, reflecting the inability to reach linear skeletal growth potential, and is therefore a key indicator of long-term well-being.

Figure 3 shows the HAZ profile by age, in the cross section of old children in the Control group at baseline (recall these are a randomly chosen set of children aged 0-60 months in the household). We see a standard U-shaped profile: early in life (at 10 months), we see HAZ scores of below -1.5, so children have poor initial conditions in terms of physical human capital accumulation relative to international standards. The HAZ scores decline further as children age, a commonly observed phenomenon in low-income settings referred to as 'growth faltering'. HAZ scores then plateau between 24 and 40 months, at which point children catch up slightly on this metric to the international benchmark. For ages 40 to 60 months, we see HAZ scores stabilize at -2.5.

Two points are of note. First, the fact that stunting is so severe in early life suggests stunting may begin *in utero* with children being born stunted. If so, children are likely exposed to chronic nutrient deprivation during pregnancy (intrauterine growth retardation). Hence the importance of the pre-natal messages and resources targeting children *in utero*. Our midline estimates, $\widehat{\gamma}_M$, are taken around the two-year mark when HAZ scores start to plateau, so we can assess whether the intervention slows down the process of growth faltering. Second, there has been a growing body of research in human biology trying to understand the causes of growth faltering in the first 24 months of life. One class of explanation relates to the returns to household resources being especially pronounced in early life. A second class of explanations emphasizes nutrition and that energy is needed for physical growth and development. We will shed light on these channels when we unpack the mechanisms driving child outcomes.

3 Child Outcomes

3.1 Fertility, Birth Spacing and Mortality

We first consider impacts of the intervention on fertility, birth spacing and child mortality. Panel A of Table 4 shows results for our core sample of women pregnant at baseline, and so immediately eligible for transfers. We find no significant ITT impacts, either at midline or endline, on fertility rates. Hence the program does not induce a lower likelihood of subsequent births among those pregnant at baseline.[17]

Panel B probes this finding to examine whether there are differences in estimated gestation of the new child (as measured at midline) or in birth spacing between the new and end child (as measured at endline).[18] For these outcomes we control for the estimated month *in utero* of the new child at baseline. The results suggest a small impact on gestation length of around two weeks (with the obvious caveat that gestation length is noisy and based on mother's self-reports). Any impact on gestation length could have been driven by women responding to CDGP messages promoting antenatal care and improved diets of mothers. An established literature also documents a relationship between maternal stress and gestation length [Currie and Slatin-Ross 2013]. A key stressor in our context is the lean season when food is scarce (and the majority of new children are born during this period). The provision of cash transfers over a sustained period might help to ease this stress, and so also help increase gestation length slightly.[19]

We also observe a significant reduction in birth spacing between the new child and end child for treated women: the magnitude of the impact is just under one month. In consequence, at endline, the end child of treated women will be slightly older than the end child in the control group. Birth intervals of less than 24 months are considered to place the child at higher risk of mortality and undernutrition, and mothers with those intervals are at a higher risk of birth complications. We see that there is a significant increase in such short birth spacing for the end child, albeit from a very low base. This will need to be borne in mind when we later consider within households spillovers on the younger sibling of the new child.

Finally, Panel C shows no treatment effects on measures of child mortality.[20]

[17] We find no robust evidence of the program impacting household composition at midline or endline, as measured by the number of individuals resident in the household in various age bins.

[18] Information on birth weight is unavailable: children are rarely weighed at birth and most are delivered at home.

[19] The magnitude of the effect amounts to more than one standard deviation in gestation lengths estimated in similar low-income contexts [Elshibly and Schmalisch 2008], including Nigeria [Okeke *et al.* 2014]. The fact the intervention begins in the final trimester of pregnancy (and transfers begin even later in pregnancy) makes any true impact on gestation less likely. The evidence on drivers of gestation in low-income settings remains scarce, partly because only noisy measures of gestational age are available in such contexts. Estimates of the effect of prenatal care on gestation length vary from zero [Evans and Lien 2005] to more than two weeks Li and Poirier [2003].

[20] In the Control group the implied mortality rate in the 0-2 year age range is : 152/1000 live births, that is higher than for Nigeria as a whole as measured by the Nigeria DHS 2013 but in line with the Northwest region being more deprived than other parts of the country.

3.2 Anthropometrics

We first consider the distribution of HAZ scores at midline and endline: Figure 4 shows there is a rightward shift of the distribution between Treated and Control new children in both time periods. Panel A of Table 5 then shows ITT impacts of the intervention on height and stunting for the new child. We see that: (i) at midline, treated children have a statistically significant increase in their HAZ score by $.20\sigma$; (ii) at the lower tail of the distribution, there is a reduced incidence of stunting of 5.22pp, corresponding to an 8% reduction; (iii) at the extreme tail of the distribution, there is a reduced incidence of extreme stunting of 4.77pp, corresponding to a 14% reduction. Most importantly, these impacts are sustained at endline, four years post-intervention and well after cash transfers have stopped being provided. We do not reject equality of the midline and endline impacts on HAZ, stunting or extreme stunting. The pattern of results remain significant at the 10% level when we adjust for multiple hypothesis testing.

These impacts are the total effect of the intervention operating through both any changes in gestation length, and at-age effects on height. Below we show the robustness of the results to flexibly controlling for age (and so only measuring any at-age impact on height). Moreover, the impacts are too large to be only driven by an effect on gestation.[21]

These effect sizes are at the upper end of documented impacts of conditional cash transfer programs in middle-income contexts, where conditionality often requires households to undertake some positive parenting practices [Sridhar and Duffield 2006]. For example, Maluccio and Flores [2004] present evidence from a conditional cash transfer program in Nicaragua and find two-year impacts on HAZ of $.13\sigma$. Macours *et al.* [2012] also present evidence from Nicaragua of a conditional cash transfer program and find two-year impacts on HAZ of $.07\sigma$, but these fade out after four years. The evidence from recent systematic reviews on the impact of cash transfers alone on child anthropometrics is even less encouraging [Manley *et al.* 2013, Caeyers *et al.* 2016], while systematic reviews of information-based interventions suggest they reduce stunting but more so in food secure populations [Bhutta *et al.* 2008, 2013].[22]

[21] To assess the plausibility of the HAZ impacts being driven only by differences in gestation length, we regress HAZ scores on age (in months) in the control group, based on old children aged 10 to 25 months at baseline. This relationship has a regression coefficient of $-.128$. Hence to generate the $.20\sigma$ increase in HAZ at midline, treated children would need to be 1.5 months younger than those in control villages. This is both implausible given the intervention began in the final trimester of pregnancy, and this impact lies outside the 95% confidence intervals estimated in Table 4.

[22] On cash transfer interventions, Sridhar and Duffield [2006] overview the impacts of a number of conditional cash transfer programs from Latin America. These generally lead to larger reductions in stunting and include evidence from Progresa in Mexico (10% reduction) and RPS in Nicaragua (5.3% reduction). they also review two unconditional cash transfer programs in Sub Saharan Africa and find neither impacts stunting. None of the programs reviewed specifically target children *in utero*. Manley *et al.* [2013] provide an updated review of 21 studies covering 17 cash transfer programs. They find the average impacts on HAZ to be positive but not statistically significant (although most programs do lead to increased food consumption). Caeyers *et al.* [2016] reiterate this view in their overview of the literature, where they state that rigorous evidence on the impacts of UCTs remain limited, with the evidence suggesting insignificant impacts on child nutrition or the impacts being limited to subgroups. On information-based interventions, Bhutta *et al.* [2008] provide a systematic review of the

To get a clearer sense of the magnitude of our estimates, we can convert the HAZ scores to unstandardized height: the midline ITT corresponds to the new child being .49cm taller, and being .62cm taller at endline. Although these are small increases – and perhaps not even noticeable to parents – they do represent economically significant population wide-impacts. Notice however that the mean difference in height for children aged two in the Control group at baseline, between the top and bottom wealth quartile is .24cm. Relative to this benchmark, the documented impacts on HAZ might in fact be noticeable to parents, and thus lead to a virtuous cycle in terms of improved child related practices.

3.2.1 Robustness

Table A5 presents robustness checks on these findings that allow for age-adjustments, so controlling for any possible impacts on gestation length (and so measuring only at-age treatment effects on anthropometrics) and accounting for non-linear impacts of the intervention across ages. We present three adjustments: (i) non-parametrically controlling for age in bins; (ii) parametrically controlling for a cubic in age; (iii) using a control-function approach to account for any endogenously driven impact on the age of the new child. For the majority of estimates, we continue to observe: (i) large and significant reductions in HAZ at midline and endline; (ii) large and significant reductions in stunting at endline; (iii) large and significant reductions in severe stunting at midline and endline.[23]

The remaining rows of Table A5 present ITT impacts on other anthropometric outcomes for the new child: we see no impact on weight-for-age Z-scores (WAZ): this is as expected given the low incidence of wasting in this population to begin with (as shown in Table 2 and consistent with data from the Nigeria DHS from 2013). The lack of impact on weight is in line with other interventions providing cash transfers early in life [Maluccio and Flores 2004, Handa *et al.* 2010, Macours *et al.* 2012, Levere *et al.* 2016].[24]

experimental and non-experimental evidence on impacts on maternal and child nutrition. They find generally weak impacts of programs promoting breast-feeding on stunting, especially when targeted to food insecure populations. In food secure populations, interventions that provided education about complementary feeding increased HAZ scores by .25 (with a 95% confidence interval of .01-.49. Bhutta *et al.* [2013] update this review, covering 110 RCTs and quasi-experiments on breast-feeding promotion in infants, and 16 RCTs and quasi-experiments on complementary feeding promotion for children aged 6-24 months. Again impacts are larger in food secure populations, although fewer studies find these translate into reductions in stunting.

[23] In specifications where age is non-parametrically controlled for, we include dummies for the following age groups (in months): 14-20, 21-27 at midline; 21-27, 28-33, 34-39, 40-45, 46-51 at endline. When using the control function approach, we use the date of interview as an instrument for age. We exploit the fact that fieldwork for each survey wave takes place over a number of months, and so children in households surveyed later are comparatively older than children surveyed earlier. The validity of the instrument is based on the assumption that the time at which households are surveyed is orthogonal to unobserved determinants of a child's physical growth. The first stage is highly predictive. We then take the first stage residuals and their square, and control for them in the second stage estimates shown (adjusting the resulting standard errors). Point estimates are similar across specifications, but for endline impacts standard errors become very large in the case of the control function estimator.

[24] Wasting reflects recent or current weight loss. As such weight-based measures are sensitive to recent illness and child feeding practices as well as seasonal variation, stunting has long been considered the more reliable indicator for identifying need in early life [WHO 1995].

Combining impacts on height and weight we find a reduction in weight-for-height Z-scores (WHZ) at midline, that fades by endline. We find no significant change in middle upper arm circumference (that is a proxy for malnourishment).

Finally, Table A6 shows anthropometric outcomes by gender of the new child: we find slightly more precisely estimated impacts for girls, although as the final two columns show, there are no significant differences by gender on any anthropometric outcome at midline. At endline only the impacts on WAZ and WHZ differ statistically by gender, while the other six dimensions do not.[25]

3.3 Health

Panel B of Table 5 shows treatment effect estimates on health-related outcomes for the new child. We first consider an index of health outcomes made up of two dummy variable components: whether the new child has not been ill in the last month, and whether the new child has not had diarrhea in the two weeks prior to the survey. This index improves significantly by $.21\sigma$ at midline, and by $.29\sigma$ at endline. Again, health impacts of the intervention are sustained well after cash transfers stop being provided.

The remaining rows show impacts on each index component: there is a reduction in illness/injury for new children of 8.53pp at midline (corresponding to a 12% fall), and this reduction improves slightly to 12pp (17%) by endline. The incidence of diarrhea among the new child also falls dramatically: at midline there is a reduction of 6.9pp (corresponding to a 18% fall), and this again rises slightly to 9pp (25%) by endline.

These kinds of health impact and their magnitude are likely to be more noticeable to parents. As such they might lead to reinforcing types of behavioral change, as we examine below when studying the mechanisms driving these new child outcomes.

The outcomes considered so far are all targeted as part of the informational messages delivered through the CDGP. In the Appendix we consider whether these improvements spillover to margins of cognitive and non-cognitive development of the new child, that are not targeted but that also have potential importance in determining lifetime welfare. Summarizing our findings from Table A7, we find muted impacts on these developmental outcomes by endline.

4 Mechanisms

Given the multifaceted nature of the CDGP, we sequence the study of mechanisms into those predominantly related to the information components of the intervention, those predominantly related to the cash component, and those reflecting a combination of both. At the end of this

[25] We have also explored heterogenous impacts on height by: (i) the low- and high-intensity information treatments (T1 vs. T2); (ii) the wife's likely control over the use of the transfers, as expressed by her at baseline; (iii) whether the marriage is polygamous or not. Along each dimension, we do not find robust evidence of heterogeneous impacts across households.

Section we use mediation analysis to shed light on the relative roles of information- and resource-related channels in mediating anthropometric outcomes for the new child. By examining how mediators vary at midline and endline, our analysis provides insights on the dynamic impacts of behavioral change caused by the intervention.

4.1 Knowledge

We first consider impacts on each parent's knowledge of pregnancy-related practices. To do so, we construct a knowledge index for each parent, built from seven questions: (i) would you advise to seek a check-up even if the baby is healthy? (ii) is colostrum good for the baby? (iii) should you breast-feed immediately? (iv) where is best place to give birth? (v) should a baby receive any other liquids on first day? (vi) should you give water to a baby if it is hot out? (vii) how long should you exclusively breast-feed for? To avoid social desirability bias in responses, these dimensions of knowledge all relate closely to the key messages provided by the CDGP on practices in ante-, peri- and post-natal periods, but this knowledge index goes beyond the literal recall of messages (that was shown earlier in Figure 2), and measures parent's ability to practically apply the knowledge in new scenarios.

The results are in the first row of Table 6 and show that: (i) women have significant increases in their knowledge index of $.95\sigma$ at midline, and $.80\sigma$ at endline; (ii) husbands have significant increases in their knowledge index of $.38\sigma$ at midline, and $.26\sigma$ at endline; (iii) the knowledge impact on husbands is smaller in magnitude than for their wife's in each period (although this is more obvious in the panel below, where we examine the components of the index individually), and this is as expected given men's weaker engagement with some information channels such as food demonstrations; (iv) for both spouses, the falls in knowledge between midline and endline are statistically significant ($p = .027$ for women, $p = .023$ for men), so that knowledge impacts are sustained at four years post-intervention, but they do fade slightly over time.

These impacts are large, partly due to low levels of knowledge at baseline, but also reflecting the quality and design of the information campaign. It is also notable that husbands' knowledge is substantially affected by this intervention. All else equal, this increases the likelihood the additionally acquired knowledge is actually acted upon in the form of better practices.

The remainder of Table 6 shows impacts on specific dimensions of knowledge. This highlights the very low levels of knowledge among the Controls on each dimension. Concretely, we observe improvements in knowledge, of women and their husbands, starting from when the new child is *in utero* (such as visiting health clinics for check ups), when the new child is born (such as giving birth in a health facility, giving the new child colostrum, breast-feeding them immediately, and giving them no other liquids on their first day), and in their first 1000 days of life (such as not giving water to children aged below six months and exclusively breast-feeding them for six months). In nearly all dimensions: (i) the magnitude of impacts is larger for women than husbands at midline and endline; (ii) there is a slight fading of knowledge between two- and four-years. The persistent impacts of improved knowledge to endline suggest there can be intertemporal benefits to younger siblings of the new child: in the Appendix, we examine whether this is so using outcomes for the end child.

4.2 Practices

Improvements in knowledge only translate to improvements in child outcomes if they are acted upon. The mapping between knowledge and practices is not assured: there is a wealth of evidence related to health behaviors suggesting limited attention, present bias and endogenous belief formation can severe any tie between knowledge and what is acted upon by individuals [Kremer and Glennerster 2012, Dorsey *et al.* 2013].

We study the issue in our context by examining impacts on the practices mothers engage in with their new child. To do so, we first construct a practices index comprised of behavior towards the new child in the ante-, peri- and post-natal periods. Panels A, B and C then show how these specific practices change with treatment, with each practice mapping to a dimension of knowledge considered earlier. We do so only at midline because by endline, these practices will be irrelevant for the new child as they turn four. We do not ask husbands to report practices as mothers are the central caregiver to the new child.

The first row in Table 7 shows that treated women significantly improve practices towards their new child: the index rises by $.88\sigma$ at midline. The Panels below reiterate the prevalence of poor practices among Controls: only 20% of mothers received antenatal care while the new child was *in utero*, only 44% put the new child to breast immediately, and only 12% exclusively breast-fed for the first six months of the new child's life. Along all five dimensions of peri-, ante- and post-natal practices, we observe statistically and economically significant improvements in mother-child practices at midline for treated women. Changes in knowledge thus do translate into changes in actual behavior towards the new child. Taken together, these changes in behavior have the potential to drive anthropometric and health outcomes for the new child during its first 1000 days of life [Kramer and Kakuma 2012].[26]

4.3 Health Behaviors

Panel D in Table 7 examines specific health-related behaviors of the mothers towards the new child. These go beyond the core messages provided by the intervention. To do so, we construct a health behaviors index comprised of two components: (i) has the child been given deworming medicine in the past six months? (ii) has the child received all basic vaccinations? We see that: (i) this index of

[26] Qualitative evidence from interviews with a subset of beneficiary households indicate widespread understanding of the practices recommended through the information component of the CDGP. Respondents were reported as embracing the suggestions after observing beneficial impacts on children [Sharp *et al.* 2018].

health behaviors significantly increases by $.17\sigma$ at midline, and increases further to $.33\sigma$ by endline with this rise over time being statistically significant ($p = .082$); (ii) the next two rows show that each individual component increases: the likelihood a child is given deworming medication in the last six months increases by 8pp (or 49%) at midline, and by 12.1pp (74%) at endline; the likelihood a child has received *all* their basic vaccinations increases threefold by endline. It remains very rare for a child to have a full set of vaccinations, and so what might be of more relevance are specific vaccination rates. Figure A3 shows ITT impacts on individual vaccinations: there are substantial increases in vaccination rates for DPT, BCG, measles, hepatitis B and yellow fever (only polio vaccinations do not increase): each rises by around 10-15pp by endline. Even putting aside all the earlier documented impacts on child anthropometrics and health, these increases in deworming and vaccination rates in early life are likely to translate to long run impacts on children and raise their lifetime welfare [Baird *et al.* 2016].[27]

The final two rows of Table 7 examine mother's health behaviors towards the new child conditional on the new child having had diarrhea in the two weeks prior to survey date.[28] We find significant improvements in the behavioral response of mothers: the likelihood they seek any advice/treatment rises by 6.9pp (9%) at midline, and by 7.6pp (10%) at endline; the likelihood the child is given oral rehydration salts (that are available from local health facilities) increases by 10pp (25%) at midline, and by 14.1pp (35%) at endline.[29]

Taken together, these changes in behavior all have potential to improve and sustain anthropometric and health outcomes for the new child. As with knowledge and practices in pregnancy, the persistence of these improved behaviors also suggests there can be intertemporal benefits to younger siblings of the new child, and contemporaneous benefits for the health of older siblings of the new child (issues we examine in the Appendix).

4.4 Dietary Diversity and Food Security

We next consider the dietary diversity of foods consumed *specifically by* the new child. We do so using two measures: (i) an overall index of the dietary diversity measuring the number of food groups the new child is fed; (ii) whether at least four food groups were consumed by the new child.

[27] Baird *et al.* [2016] present experimental estimates on the long run impacts of a school-based deworming program. They find that ten years after deworming treatment, men who were eligible as boys stay enrolled for more years of primary school, work 17% more hours each week, spend more time in nonagricultural self-employment, and are more likely to hold manufacturing jobs. Women who were in treatment schools as girls are approximately one quarter more likely to have attended secondary school, halving the gender gap. They reallocate time from traditional agriculture into cash crops and nonagricultural self-employment. They estimate an internal rate of return to deworming of 32%.

[28] As the intervention impacted the incidence of diarrhoea, children in this sample may not be exactly the same in Treatment and Control groups.

[29] By endline we also find significant improvements in the likelihood that soap is at the place for hand washing in the household, and in the quality of toilet facilities (in line with messages provided on sanitation). At the same time we find no evidence of households gaining access to improved water sources, that is as expected given individual households can do little to drive forward such infrastructure improvements.

Both measures are obtained from a 24-hour food recall module administered to the new child's mother or main carer, at midline and endline. Each meal consumed by the new child in the day before the interview from waking up to bedtime is recorded, with ingredients of each meal being coded into seven food group categories.[30]

The results are in Panel A of Table 8. We see that: (i) the dietary diversity index for the new child rises by .36 (or 11%) at midline and this improvement is sustained at endline. The likelihood that at least four food groups are consumed rises by 10.7pp (23%) at midline, and by 12.7pp (27%) at endline. The dietary recall data allows us to examine the exact food groups consumed by the new child. This breakdown is shown in the first set of Columns in Table A8. The food groups driving increased dietary diversity are dairy products, flesh food and eggs, and other fruit/vegetables. The fact that two of these relate to produce derived from livestock is important to bear in mind, as we consider other mechanisms more closely linked to the cash component of the CDGP, such as impacts on labor supply and investments into business assets.

We probe the data further to understand whether changes in food diversity, as measured by 24-hour recall, reflect more sustained dietary changes over the course of the year. To examine this, Panel B of Table 8 examines the food security households report in the 30 days prior to midline and endline surveys. We do so in an economic environment where there is a lean season for agriculture and food production: in the control group, 16% of households report not having had enough food to eat in the month prior to the midline survey. We see significant improvements in food security: food insecurity falls by 4.4pp (27%) and by midline accelerates to a 9.3pp (57%) reduction, the difference between the two being significant ($p = .020$). By endline there are also statistically and economically significant reductions in the share of households reporting having gone the whole day and night without eating, and ever going to bed hungry – with the incidence of the latter almost being eliminated altogether among treated households.

All endline ITT estimates are robust to adjustments for multiple hypothesis testing.

We find improvements in food security both at endline (which, from Figure 1, takes place during the lean season that runs from March to October) and at midline (that takes place outside the lean season). Given the high levels of background uncertainty that households face in this environment, Table 9 shows how food security is impacted *by season*. We see that: (i) throughout the year there are significant improvements in food security, and these are most marked during the lean season (Damuna, that runs from June to October); (ii) these improvements become larger at endline than midline; (iii) multiple reasons are provided for why food security has improved

[30] To map from meals to food groups, our enumerators proceeded as follow. They first listed the dishes consumed by the new child in the 24-hour recall module (excluding drinks – these were captured separately in the liquids recall module), and then coded up the individual ingredients used in each dish as reported by caregivers. Although in theory this ingredient list can be very long, in practice the dishes consumed did not vary a lot. At a final stage, the ingredient were then mapped to food groups. These food groups are: (i) grains, roots and tubers; (ii) legumes and nuts; (iii) dairy products; (iv) flesh foods; (v) eggs; (vi) vitamin-A rich fruits and vegetables; (vii) other fruits and vegetables.

including having more resources; (iv) this all leads to households being less reliant others in informal risk sharing networks to cope with food shortages, or having to engage in more extreme forms of coping strategy – such as selling livestock or just consuming less – that are not in their own long term interest.

Together, the results thus highlight not only improved nutrition on a given day for the new child, but also improved availably of food for treated households both during the lean season and at other times. Both mechanisms can potentially drive the positive impacts on new child outcomes documented earlier. While dietary diversity can be driven by information provision alone, we note that food security improves even more at endline than midline. This is remarkable because the endline occurs well after these households are in receipt of cash transfers from the program itself, suggesting there might be long lasting impacts on the resources available to treated households, even after cash transfers end.

We thus next examine mechanisms more closely related to the provision of cash transfers.

4.5 Labor Activities

There are two substantive reasons why the cash transfers provided can impact child outcomes beyond any direct effect on health-related behaviors or food purchases. First, the value of the cash transfer – US$22 per month – was calibrated by our intervention partners to correspond to the cost of a diverse household diet. However, we already noted that at baseline Control households spend $85 per month on food suggesting a potential crowd out of resources for other uses, and Figure A1D suggested that households have the possibility to improve nutritional intake without changing food expenditures. Second, the fact that households are aware that transfers will be provided each month until the child is 24 months old, provides women with a more stable flow of resources than is available from most labor activities in these rural economies. The magnitude and certainty of transfers opens up the possibility that they are used for both investment and consumption outcomes.

The results are in Table 10 and can be summarized as follows: there are marked and permanent changes in the labor supply of women, in business investments made by women, with little change in the labor activities or business investments of men. This leads to long run earnings increases for women, amounting to large *sustained* increases in resources available to households in the period after cash transfers are being received as part of the intervention.

We break down this chain of analysis as follows. Panel A focuses on the labor activities individuals are engaged in, so the extensive margin of labor supply. In this setting women's labor force participation rates are high to begin with (74% at baseline in the Control group). For treated women this rises by 6pp by midline (despite these women being pregnant at baseline and so unable to work continuously between baseline and midline), and by 11pp by endline. By endline, women also become more likely to engage in multiple activities, and there is a significant increase in the number of days per week spent in their highest earning activity, so on the intensive margin of labor supply. This is all consistent with treated women being able to generate more diverse earning streams by endline.[31]

The right hand side of Panel A shows much smaller impacts on husband's labor supply.

Panel B focuses on the types of labor activity engaged in: recall that at baseline the most common activities for women are being self-employed running a small-scale business, such as livestock rearing or petty trading. We see significant increases in self-employment and petty trading activities at midline, with impacts increasing in magnitude at endline.[32]

The right hand side of Panel B shows no corresponding impact on the labor activities of husbands: they are mostly engaged in farming their own land and the incidence of this does not change post-intervention.[33]

Given the labor activities women engage in, we next focus on two types of business investment: expenditures on business inputs into the woman's own business and livestock ownership. We see both types of productive investment being undertaken after cash transfers have been provided.

On business inputs, these increase significantly by $21/month at endline, long after cash transfers were last provided. We see no corresponding increase in expenditures on inputs for the husband's business, again suggesting there are no large resource transfers across spouses.[34]

Regarding livestock ownership, women's ownership of any animal increases by 6.6pp (8%) at midline, and by 12.1pp (15%) at endline. These impacts are statistically different of each other ($p = .016$). Livestock ownership is critical in this economic environment because: (i) it generates earnings for women from the sale of animal produce such as milk and eggs; (ii) it produces an earnings stream all year round thus reducing the volatility of earnings women are subject to; (iii) animal produce can also be consumed at home, and this maps closely to the documented impacts on dietary diversity of the new child in Table 8. The increased dietary diversity of foods given to the new child was driven by the increased consumption of dairy products, flesh food and eggs. Such protein-rich foods have been argued to, if consumed at critical ages early in life, drive physical growth and neurological development and potentially slow down the pattern of growth faltering seen in HAZ rates in low-income settings [Dewey and Adu-Afarwuah 2008, Headey *et al.* 2018].[35]

[31] The extensive margin responses might reflect that the resources enable women to overcome fixed costs of working, such as being able to travel to work, or pay others to look after young children.

[32] These results are supported by the parallel qualitative workstream that interviewed beneficiaries: this shows women invested into small-scale home-based activities such as petty trade, food processing and sale, small livestock rearing, and services to other women (such as hairdressing or pounding grain) [Sharp *et al.* 2018].

[33] We have also used the data to probe further on impacts on agricultural inputs and crop cultivation. We find muted impacts on husband's expenditures on seeds and fertilizer, with a 25% increase in pesticide expenditures by endline. On crop cultivation, we find no significant impacts – at midline or endline – on crop types cultivated on husband's land (the majority of which remain grains, tubers and roots).

[34] This in contrast to recent findings from other settings that show how women's capital can often be invested into their husband's micro-enterprise [Bernhardt *et al.* 2019]. In our context, as stressed from the outset, women retain bargaining power over income streams they earn and this is reflected in these measured labor and investment responses by gender.

[35] Headey *et al.* [2018] describe how cow's milk (an important source of amino acids, calcium, iron, and vitamin

Given the potential importance of the links between livestock, earnings and nutrition, we probe this finding in two dimensions.

We first detail livestock ownership of households, and women themselves. Table A9 shows: (i) increases in household ownership of livestock are largely driven by livestock owned by treated women (and not another household member); (ii) the ITT point estimate on owning any given animal is always higher at endline than midline; (iii) the main types of livestock women become more likely to own are goats, chickens (at midline and endline), and by endline, sheep, donkeys and calves.[36]

Second, we examine whether the cash transfers provided by the program are sufficient to plausibly allow women to purchase these kinds of lumpy asset. Table A10 shows mean and median unit prices of livestock from two survey questions fielded to Control households at baseline: (i) prices paid to purchase an animal; (ii) revenues from sales of such animals. Obviously, these prices are based on select samples, and do not account for livestock quality. However, they provide an indication of the plausibility of the findings on livestock ownership. The highest median unit price for any livestock type (male sheep) is $121 based on purchases and $201 based on sales. These values correspond to between six and ten months worth of CDGP transfers: recall that these transfers are valued at $22 per month, and that by midline, the cumulative value of transfers received by women pregnant at baseline is $470. This all suggests: (i) by midline it is feasible for investment into livestock to be sunk; (ii) this would still leave the majority cumulative value of transfers received available for other uses, including other business investments, consumption and savings accumulation (as we examine below).[37]

Panel D of Table 10 combines all the information on changes in labor activity to construct a (noisy) measure of total monthly earnings from all forms of employment, for each spouse: we see at midline women's earnings increase by $19.4 (corresponding to 20%), and the magnitude of these earnings are exactly sustained at endline. In line with all the earlier results, we see no statistically significant impacts on earnings of husbands.[38]

B-12) stimulates the secretion of insulin-like growth factor I (IGF-I), the hormone that stimulates bone and tissue growth; eggs are an excellent source of choline, that is needed for the synthesis of phosphatidycholines, a process relevant for bone formation and cell membrane formation.

[36] We have also examined the number of livestock owned (where we asked this question for larger animals, but not for poultry). We find that by endline there are significant increases in the number of calves and sheep owned by women. This suggests the impacts on livestock are driven both by women investing in livestock for the first time, and by others expanding existing herds.

[37] Credit market imperfections likely restricted the ability of households to borrow to purchase livestock pre-intervention. However, we also note that household savings at baseline among the Control group are valued at $272. This means *ex ante* households were able to purchase such livestock even absent CDGP transfers if they were willing to use half their stock of savings. However, given the volatility of the economic environment, households likely have a strong precautionary savings motive.

[38] The increased earnings are generated through changes on the extensive and intensive margins of labor supply, as well as returns to business investments. However, another potential channel could be that as women's nutrition improves, they become more productive in existing activities. We lack detailed data on labor productivity, although in the Appendix we document largely null impacts on the health of treated mothers in terms of their anthropometrics.

4.6 Expenditures, Savings and Borrowing

Having described impacts on labor activities, investment and earnings, we now complete the household budgeting exercise by examining impacts on expenditures, savings and borrowing. Food expenditures are calculated based on a seven-day recall, by food group. These map to the same food groups considered in the dietary diversity measure. Expenditures thus relate to *flows* at midline and endline. In contrast, savings/borrowing relate to *stocks* accumulated between surveys. The results are in Table 11.

Panel A shows ITT impacts on expenditures. Starting with food purchases, we see that monthly food expenditures rise by \$21 (25%) at midline, and this increase is largely sustained at endline where they are \$16 higher than the Control group. We can break down food expenditures by food groups. These results are shown on the right hand panels in Table A8, thus facilitating comparison to changes in food consumption as show on the left hand panels of the same table. We see that: (i) there are significant increases in expenditures at midline and endline on dairy products, and other items (including sugary items and drinks); (ii) by endline, there are increased expenditures on legumes and nuts, and oil, butter and other condiments; (iii) no food group has a decline in expenditure over time.[39]

Figure 5 pulls together the various strands of impact on investment into livestock, food consumption and food expenditures. The left hand Figure shows percentage impacts at endline on women's livestock ownership where we classify animals in terms of produce (commonly eaten, egg producing and milk producing). The right hand side figure shows for each food group, the percentage impacts at endline on dietary diversity for the new child, and household expenditures.

This reconfirms that increases in livestock types map closely to compositional changes in dietary diversity: the largest percentage impacts on dietary consumption of the new child are for flesh food and eggs and dairy products, that are all sources of animal protein. For both animal products, the percentage increases in consumption are far larger than the corresponding increases in expenditures. This reinforces the notion that these food groups are produced at home, through investment in livestock ownership that is financed by the cash component of the program.

Returning to Table 11, we combine food and non-food expenditures to estimate that total expenditures rise by \$40.4/month at midline, falling back to \$25.2 by endline (but still being sustained well after cash transfers have been disbursed).[40] The magnitude of this increase at midline corresponds closely to the sum of additional resources available to the household via CDGP transfers (\$22) and the increase in women's earnings shown in Table 10 (\$19). By endline,

[39] We have also estimated quantile treatment effects on monthly food expenditures: we find no robust evidence of a difference in impacts across expenditures deciles. The same applies to monthly total expenditures.

[40] Non-food expenditure is obtained combining the following sources: (i) a 7-day expenditure recall of consumables (e.g. matches, fuel); (ii) a 30-day recall of other items (e.g. toiletries, utensils, household items, health expenditure); (iii) a 12-month recall of major expenses (e.g. school fees, ceremony costs, remittances); (iv) expenditure on durables using a 12-month recall of expenditure on assets the household owns (e.g. TV set, wheelbarrow, mattress). The top 1% of total expenditure amounts are trimmed.

the increase in total expenditures closely matches the increase in women's earnings ($19).

The share of total expenditure on food does not rise significantly at midline but does so by 2.2pp by endline. The fact that food shares do not decline as overall expenditure increases also suggests that there may have been a shift in the household Engel curves for food. This could be due to either a change in preferences of the household (say driven by the knowledge impacts of the program), or changes in women's bargaining power driven by the transfers provided to them. We cannot examine this directly because we only collected information on bargaining power at baseline. However the baseline data reveal that while women have relatively weak decision making rights over many dimensions of household decision making, they retain autonomy in how to spend additional resources they bring into the household.[41]

In line with this, at midline we asked who usually decides how to spend the CDGP transfer: nearly 75% of women, and 75% of husbands, reported the wife alone decided. Women thus appear to have major control over the use of the transfer, and this may point to some degree of non-cooperative bargaining in these households [Browning *et al.* 2010]. This all fits firmly with the earlier results suggesting that cash transfers to women do *not* leak away to be invested in the economic activities of their husbands.[42]

Panel C the examines the stock of savings and borrowings of the household (Table A11 provides more disaggregated information on impacts along these dimensions). We see that by endline there is a significant rise in household savings of $55.6, and a significant reduction in borrowing of $20. Both changes help households build resilience to shocks, that as documented in Table 1, is important in this economic environment given the frequency of aggregate shocks.[43]

4.7 Net Resources and Extreme Poverty

We conclude our budgeting exercise by drawing together all changes in resources inflows and outflows to derive an implied change in the net resources available to the household at midline and

[41] That baseline data reveal that women have relatively weak decision making rights in relation to major household purchases, which food to grow, what food to buy: along all three dimensions, the majority of women report their husband decides alone (or in consultation with them), and very few women ever make these decisions alone. This pattern is confirmed in both interviews to the wife and their husband. However, we also asked a series of vignette questions at baseline on who would have decision making rights over any new flow of resources that the wife generated. In these scenarios: (i) the majority of women reported they would decide alone how to spend the new resources; (ii) this was so irrespective of how the additional resources were generated (either through labor earnings, or as a gift to the wife); (iii) husband's reports were near identical to their wives in all cases.

[42] When asked at midline to report what most of the cash transfer was used for, the most frequent responses of women were food for the household (64%) and food for children (24%). Husbands provided very similar reports.

[43] Panel A of Table A11 shows the positive impacts on savings exist on the extensive and intensive margins: the share of households able to save at all rises by 6.7pp at endline (corresponding to a 16% increase over baseline). Panel B shows that on borrowing, the reduction in borrowing occurs at the extensive margin with treated households being 7.5pp (32%) less likely to have any member borrowing at endline. On a crude proxy of borrowing constraints (whether any household member failed to borrow funds when the desired to do so) we see little impact of the intervention, that is in line with expectations. Finally, Panel C shows that there are no significant changes in household lending at endline on either the extensive margin or the amount of funds lent to others.

endline: this includes the exogenous receipt of cash transfers from CDGP for treated households (up to midline), and endogenous changes in earnings arising because of the intervention. The imputed value of net resources is calculated as spousal earnings + savings − borrowing + CDGP transfer, where each element is computed as a monthly flow at survey date. As saving and borrowing are measured as stocks, we convert these into monthly flows assuming they accumulate at a constant rate between survey waves.

Panel C in Table 11 shows the result: we see an increase in net resources available to the household. More importantly, the magnitude of the increase is $49 at midline, so more than double the value of the cash transfer itself. This suggests the multifaceted CDGP program induces large behavioral responses of household members that endogenously generate increased resources to the household. This increase in net resources is sustained at endline because the loss of transfers from CDGP is offset by an increase in net savings. The marginal propensity to consume out of these net resources is .43 (.33) at midline (endline) if only food expenditures are considered. Intertemporal consumption smoothing suggests households are more likely to consume out of these transfers if they think they are likely to persist, but these marginal propensities are lower than estimates from some conditional cash transfer programs as well as for unconditional cash transfers.[44]

The fact that the estimated elasticity of food consumption is far less than one suggests these households do not face a nutrition poverty trap [Dasgupta 1997]: improved labor productivity is not what drives the labor supply responses of women documented earlier.[45]

The final row in Table 11 considers the impact on household poverty, using the progress out of poverty index (PPI). For each household, the PPI is calculated through a scorecard and its value, ranging from 0 to 100, represents the likelihood a household is above the global extreme poverty line ($1.90 per person per day), so increases in the index represent reductions in poverty. We see that by endline, there is a 5% reduction in extreme poverty among households. This is a dramatic and sustained reduction achieved by an intervention predominantly designed to improve early life nutrition and provide resources for the first 1000 days of life of one specific child.[46]

In the Appendix we extend our analysis to consider additional outcomes: (i) maternal health (ii) within household spillovers to siblings of the new child. To summarize our findings: (i) we find

[44] For example, Angelucci *et al.* [2018] document that among *Progresa* beneficiaries in rural Mexico, the marginal propensity to consume out of their own resource transfer is .69. Almas *et al.* [2019] use an RCT providing unconditional cash transfers to document the elasticity of food expenditures to be .78, that is higher than most non-experimental estimates.

[45] We can also re-estimate this elasticity based on specific food groups, using the expenditure impacts on the right hand side of Table A8. We find no food group has an expenditure elasticity close to one, although we cannot altogether rule out a protein-related nutrition trap because the livestock investment channel creates a wedge in the calculated protein elasticity.

[46] These impacts compare favorably with other anti-poverty interventions. Baird *et al.* [2019] document that in low-income settings, there remains limited evidence on sustained long run impacts of cash transfers. Bandiera *et al.* [2017] evaluate the long run impacts of a livestock asset transfer program in Bangladesh: an intervention explicitly designed to reduce poverty (and where take up of the livestock transfers was close to 100%). They find poverty rates fell by 8.4pp four years post intervention.

no robust impacts on mothers health; (ii) we find contemporaneous and intertemporal spillovers for the health of older and younger siblings of the new child, driven mostly by sustained changes in knowledge, practices and health behaviors of mothers towards their children (although there is also some suggestion that height outcomes for older siblings worsen somewhat). The positive health effects for older siblings occur despite them lying outside the 1000-days of life window that policy often focuses on [Victora *et al.* 2010].

4.8 Mediation Analysis

We use mediation analysis to understand the relative importance of the above mechanisms, following the approach of Gelbach [2016]. The basic intuition is that the treatment effect of intervention T on outcome Y can be decomposed as operating through as set of mediators, m_j:

$$\frac{dY}{dT} = \sum_{j=1}^{k} \frac{\partial Y}{\partial m_j} \frac{\partial m_j}{\partial T} + R, \tag{2}$$

where R is the part of the treatment effect which cannot be attributed to any observed mediator. The method has the advantage that it is invariant to the order in which mediators are considered. It does not represent causal mediation except under strong assumptions [Imai *et al.* 2010].[47] However, it remains a useful exercise because: (i) it starts to unpack the relative merits of each component of the multifaceted intervention; (ii) it helps contrast the mediating impacts of channels operating through women and their husbands, and at midline and endline. Both aspects are informative for thinking through the design and targeting of interventions designed to foster human capital accumulation in early life in such economic environments.

We present the mediation analysis for the HAZ of the new child. To capture the range of information- and resource-related mechanisms documented, we use the following set of mediators: the woman's knowledge index, the husband's knowledge index, the health behaviors index, monthly food expenditures, a dummy for whether the woman works, and a dummy for whether she owns livestock. We conduct the analysis at midline and endline separately. Given the persistence in HAZ scores, at endline we allow the midline outcome to be an additional mediator.

The mediation analysis is shown in Figure 6, where Table A12 shows the underlying estimates used.[48] Focusing first on the top bar, the width of the bar shows the overall ITT impact at mid-

[47] An underlying assumption is that the mediators do not impact each other. In our context this is hard to maintain because practices related to exclusive breast-feeding, health, diet and sanitation all likely interact with each other. By ignoring such interactions, we leave more of the overall treatment effect unexplained (R).

[48] To implement the approach, consider two ITT estimators: (i) an *unconditional ITT*, i.e. $\hat{\beta}_1^U$ from the population linear model $Y = \alpha + \beta_1^U T + u$; (ii) a *conditional ITT*, i.e. $\hat{\beta}_1^C$ from the model $Y = \alpha + \beta_1^C T + M\beta_2 + u$, where M is a vector of k post-intervention mediators. Gelbach [2016] decomposes the difference between these estimators using a simple omitted variable formula: $\delta = \beta_1^U - \beta_1^C = \gamma_1^1 \beta_2^1 + \cdots + \gamma_1^k \beta_2^k = \sum_{j=1}^{k} \gamma_1^j \beta_2^j$, where $\delta^j = \gamma_1^j \beta_2^j$ is the component due to each mediator, and γ_1^j is the coefficient of the linear regression of M_k on T. These components are shown in Table A12 for each outcome.

line of the treatment on the HAZ ($\frac{dY}{dT}$): $.17\sigma$. This bar is then decomposed into the estimated contributions of each mediator. We show predominantly information-related mediators in shades of blue (woman's knowledge index, husband's knowledge index, health behaviors index), and predominantly resource-related mediators are in shades of yellow/brown (monthly food expenditures, whether the woman works, whether she owns any livestock). The unexplained part of the ITT effect is shaded in gray (R).

For the new child HAZ at midline: (i) the single most important mediator is mother's knowledge, accounting for 60% of the HAZ ITT at midline; (ii) information-related mediators are more important in explaining the HAZ ITT overall; (iii) the mother's knowledge index is a more important mediator than the father's knowledge mediator; (iv) the most important resource-related mediator is if the women owns any livestock, that accounts for 10% of the HAZ ITT at midline.

The lower panel then shows the decomposition of the total ITT effect on the new child's HAZ at endline, where we allow the HAZ at midline to be an additional mediator. We see that: (i) the HAZ at midline is the single most important mediator, confirming high persistence in HAZ over time; (ii) over and above this persistence, we see that resource-related mediators play a far larger role in mediating the HAZ at endline: food expenditures and whether the woman owns any livestock are the two most important mediators.

These results provide insights on the dynamics of program impacts: in the first two years of life, information channels are key to driving the HAZ of the new child. Conditional on the midline HAZ, the impact of treatment on HAZ at the end of four years is then predominantly determined by resource based channels. These results help explain why the provision of resources alone is unlikely to move the needle on early life child outcomes – as found in meta-analyses of the impact of cash transfers on child anthropometrics [Manley *et al.* 2013, Caeyers *et al.* 2016]. Information is key and complementary to cash, as being suggested by a nascent literature examining multifaceted interventions to drive human capital accumulation in early life, that has so far been based in contexts outside of Sub-Saharan Africa [Levere *et al.* 2016, Ahmed *et al.* 2019].[49]

Finally, the results speak directly to the growing body of research in human biology trying to understand the causes of growth faltering in the first 24 months of life: our results suggest these are more driven by information-nutrition channels than resource-related channels *per se*, with the latter becoming more important for physical growth after the first 1000 days of life.

[49] This also has implications for poverty alleviation policies such as livestock asset transfers, that have been shown to permanently life households out of poverty in South Asia and Latin America [Banerjee *et al.* 2015, Bandiera *et al.* 2017]. Our evidence suggests that in terms of child outcomes in early life, giving livestock is not sufficient. They need to be combined with the provision of context specific information on child-related practices.

5 IRR

We now derive the cost effectiveness of the CDGP and provide an *indicative* internal rate of return to the intervention. We assume the social planner has a 5% discount rate, and present the breakdown of results in Table 12.

Panel A describes program costs. We assume: (i) the per beneficiary cost to the social planner of administering cash transfers are 10% of the actual per beneficiary value of the transfers; (ii) the organization of community volunteers and other logistics to deliver the information messages amounts to a further 10% of the per beneficiary value of cash transfers.

Following the discussion in Dhaliwal *et al.* [2012], we consider two alternatives to account for cash transfers from the social planner's perspective: (i) we can view them as being a *pure redistribution* of resources from the planner to beneficiaries, and so they involve zero net cost to society; (ii) at the other extreme, we can view them as a *pure cost* solely borne by the planner, with no measured benefits to households. We focus our discussion on the first scenario, as shown in Columns 1 to 4 of Table 12, and then return to repeat the analysis under the second scenario (Columns 5 to 8).[50]

On benefits, in Column 1 of Table 12 we ignore any gains to children and only place a monetary benefit on the *net* resource flow increase to households arising from endogenous responses to the intervention. As was described earlier, these combine impacts through increased earnings (because of women's endogenous labor supply responses) and net savings accumulated. We assume these net resource flows last five years, and we use our ITT estimates on monthly net resources at midline and endline to calibrate this five year flow of benefits. As shown in Panel B, the NPV of these gains are high because they are large relative to the size of transfers, occur soon after the intervention starts, and are assumed to last five years. In consequence, the gains-cost ratio is over 12 and the IRR is over 200%.

In Columns 2 to 4 we ignore these gains in net resources to households and focus entirely on gains arising through lifetime earnings for the new child from the increase in their HAZ caused by the intervention. To do so, we exploit anthropometric-earnings profiles estimated in the longitudinal analysis of Hoddinott *et al.* [2013]: they suggest a 1σ increase in HAZ at age 24 months leads to a 4% (9%) increase in annual earnings for men (women).[51] We combine these with our

[50] We thus ignore any deadweight loss of taxation that would be incurred in order the raise the intervention cost. As Dhaliwal *et al.* [2012] also state, we exclude them because there are no reliable estimates of the magnitude of such distortions in this context.

[51] Hoddinott *et al.* [2013] almost uniquely can estimate such anthropometric-earnings profiles: they do using data from 1338 Guatemalan adults aged 25-42 in 2002, who were studied as children in 1969-77 as part of a community-randomized food-supplementation trial. Thomas and Strauss [1997] report that in Brazil, a 1% increase in height leads to a 2.4% increase in adult male earnings in a regression of log hourly wages on height and completed grades of schooling, controlling for selectivity into employment. Grantham-McGregor *et al.* [2007] document that short height among adults (a result of childhood stunting) is associated with reduced adult earnings in 55 countries. Being stunted in early childhood is associated with lower adult wages at both the individual [Hoddinott *et al.* 2008] and country level [Fink *et al.* 2016]. We focus on height rather than stunting to reflect that there economic returns

midline ITT estimates for males and females of $.14\sigma$ and $.24\sigma$ respectively to calculate percentage earnings impacts. We estimate a life-cycle profile of earnings in our sample by gender, and assume the percentage impact of HAZ on earnings is constant over the life cycle.[52]

Column 2 (3) shows cost effectiveness and the IRR for boys (girls), and Column 4 shows results for the average child. As these earnings gains start to accrue once the child is in the labor market (from age 16) but the costs are born up front (starting when the child is *in utero*) the NPV of these gains are small, even though this flow of benefits lasts many years ($60 - 16 = 44$). The IRR of the intervention for boys is 5.8%, and 3.8% for girls, with an IRR for the average child of 4.9%.

Clearly, any attempt to calculate the cost effectiveness of early life interventions is heroic because such programs impact multiple outcomes. Hence any such calculation is bound to miss many potential benefits (as well as long term costs [Alderman *et al.* 2016]), such as those arising from better nutrition, less sickness, and increased rates of de-worming and vaccinations. In Panel E we thus provide indicative estimates of what a full cost-benefit analysis might look like once we factor in additional flows of pre-labor market benefits to children from the intervention, from age 2 to 16. To provide plausible benchmarks for how large such unmeasured benefits might be, we note that per capita food consumption at baseline in controls is $11/month. We then recalculate the IRR in Columns 2 to 4 assuming these additional non-measured annual benefits are equivalent to 1, 2, 6 or 12 months of per capita consumption to the child, for each year from age 2 to 16. As shown in Columns 2 to 4 of Panel E, the IRR to the program for the average child lies between 8 and 63% across scenarios. This approach then provides scenarios with comparable estimates of the IRR to early life interventions as in high-income settings where a fuller range of benefits can actually be accurately monetized [Heckman *et al.* 2010].[53]

In Columns 5 to 8 we repeat the analysis under the scenario that the cost of cash transfers is entirely borne by the social planner (and so they generate no gains to the household or new child). As expected, the corresponding gain/cost ratios in Row C are far lower, as are the baseline IRR estimates. Hence, in this extreme accounting scenario where cash transfers represent pure costs, it becomes essential to factor in additional benefit flows to the new child over childhood in order for the social planner to find it worthwhile to invest in such an intervention. Doing so, in Panel E we find that if we value these unmeasured annual gains as equivalent to the annual value of per capita consumption, then the IRR for the average child beneficiary rises to 12%.

to stature through the distribution of height, and not just for those crossing a particular threshold.

[52] To estimate life-cycle earnings we take the cross section of women and husbands at baseline and run an OLS regression of earnings on 10-year age dummies (16-25, 26-35 etc.).

[53] Heckman *et al.* [2010] calculate the IRR to the Perry Preschool Program, an early childhood education program conducted at the Perry Elementary School in Ypsilanti, Michigan, during the early 1960s. Perry researchers collected administrative data on school records, police and court records, and on welfare participation. Their IRR calculation uses this data and accounts for compromises in the randomization protocol, the lack of program data past age 40, missing data for participants before age 40, and valuing non-market outcomes such as crime. They estimate the overall social rate of return to the program to be between 7% and 10%.

6 Conclusions

In 2015, 159 million children were estimated to be chronically malnourished, as measured by stunting or low height-for-age, so at risk of failing to achieve their genetic potential for physical and cognitive development. Childhood stunting has lifelong consequences for health, human capital and poverty alleviation [Kakietek *et al.* 2017]. By some estimates, eradicating stunting would generate hundreds of billions of dollars in benefits over the productive lives of beneficiaries in low- and middle-income countries. As such, understanding which interventions create persistent gains to human capital from early life and are cost effective lies at the very top of the development policy agenda.

We have studied the longer-run impacts of a large-scale multifaceted intervention designed to improve early life nutrition and well-being in a population with high rates of child malnourishment. The impacts of the intervention are remarkable in many dimensions. On early life outcomes, we find large and sustained improvements in human capital accumulation among children: there are notable reductions in rates of stunting, and improved health outcomes. Yet the intervention has impacts beyond the targeted child, as it transforms the economic lives of women: the intervention boosts womens labor supply, and allows them to expand self-employment activities through investing in complementary livestock assets. We see marked increases in dietary diversity (driven by the consumption of animal produce), food consumption and net savings. Overall, the combined exogenous receipt of cash transfers and endogenous female labor supply responses imply the net resources available to households increase by more than double the value of the cash transfer itself. These increases in resources are sustained long after cash transfers stop being provided, and the steady flow of earnings generated through livestock rearing helps households build resilience to shocks throughout the year including during the lean season when food is typically scarce.

Taken together out findings show the promise of a cost effective, sustainable and scalable early childhood interventions in even the most challenging economic environments.

Our future research agenda is structured as follows.

First, there is a need to understand whether the intervention continues to produce long term change in the human capital and well-being of beneficiaries. We aim to engage in future data collection with these children and households, to measure whether new dimensions of human capital accumulation, related to cognitive and non-cognitive traits, start to emerge. This question is especially pressing given the program is designed to be scalable: it is implemented in an economic environment with low state-capacity, extreme poverty and high degrees of household vulnerability. It does so by leveraging off existing resources, namely using local health facilities and hiring community volunteers. It is an intervention that could realistically be scaled-up in other parts of Nigeria, or transported to other fragile regions where almost children face significant risks of never being able to develop to their full potential, because of early exposure to severe malnourishment and extreme poverty.

Second, this evaluation has focused primarily on the 3600 sampled women identified as pregnant at baseline and so immediately eligible for cash transfers. However, we purposefully surveyed an additional 1700 women, that were not pregnant at baseline but expected to become pregnant over the course of the evaluation. In ongoing work we exploit this sample to understand endogenous responses in fertility to the provision of high-valued cash transfers to pregnant women, and endogenous responses in fertility to the announcement that the program will end. Both margins of response are vital to understand, especially given the increased roll out of unconditional cash transfer programs, often targeting women, throughout the developing world.

Finally, we have conducted our quantitative evaluation in close collaboration with a parallel stream of qualitative analysis, based on a subset of our surveyed households [Sharp *et al.* 2018]. While we have referred to the consistency of key findings across workstreams, there remain many hypotheses raised by the qualitative analysis that are of economic interest. One example is the suggestion in the qualitative work is the key role that universality of the program plays: recall that to be eligible, women have to be confirmed as being pregnant, but there is no poverty threshold at which they become eligible. The qualitative work suggests this is key in driving behavioral change, as older and wealthier women act as role models. We plan to explore this and other hypotheses raised in the qualitative workstream more systematically in future quantitative work. This helps pinpoint complementarities in these approaches, and suggests how to efficiently promote their dual use in future program evaluations.[54]

A Appendix

A.1 Cognitive and Non-cognitive Development

We consider whether the improvements in height and health spillover to margins of cognitive and non-cognitive development of the new child, that are not targeted but that also have potential importance in determining lifetime welfare. We measure development of communication and gross motor skills using modules adapted from the Ages and Stages Questionnaire [ASQ-3, Squires and Bricker 2009]. At endline we added a modified ASQ module measuring personal-social skills of the new child.

These modules assess a child's development by asking his/her caretaker whether the child is able to perform a number of specific tasks. There are six age-specific tasks (in windows of 2-3 months) asked about along each domain. For example, for motor skills, the caretaker of a child aged 19-20 months is asked, "Does the child run fairly well, stopping himself/herself without bumping into things or falling?" A child then receives zero points if child does not perform the

[54] Bergman *et al.* [2019] provide a recent example in economics of the benefits of blending analysis between quantitative and qualitative workstreams: they do so in the context of using a randomized control trial to study the impacts of housing vouchers on social mobility among recipients in Seattle and King County.

task yet; five points if the child performs it "sometimes"; ten points if he/she does it habitually. We convert aggregate scores on each domain to Z-scores based on international norms. We also report impacts on the likelihood of being below specific thresholds (say for low communication skills), below which children (in richer countries) should typically be referred to a developmental nurse or psychologist for further assessment. In the absence of locally validated thresholds, we use the thresholds from the reference western population.

Panel A of Table A7 shows the results. At midline there are significant impacts on the communication skills of new children, with a 7% reduction in those classified as having low communication skills. However, these impacts fade out over time. We find no evidence the program impacts motor skills or personal-social skills.[55]

This is despite the fact that as shown in Panel B, the time mothers allocate towards the new child does increase (that was only measured at midline). We find a shift away from mothers reporting spending less than two hours playing with their child to an increase in the share reporting spending more than five hours playing with the child. If key investments into children are time intensive, they seem to feed through into anthropometric and health outcomes, but not domains of child development.[56]

A.2 Maternal Health

Maternal health is critical to infant survival and child development. Some of the key messages provided by the intervention relate to mothers maintaining their nutritional status, in recognition of the fact that energy and nutrition needs increase during pregnancy and lactation. Moreover, improved food expenditures, dietary diversity and food security through seasons might raise women's labor productivity. In turn this can drive the labor supply responses documented earlier.

Table A13 shows impacts on maternal health using anthropometric outcomes. At midline we find no robust evidence of changes in any health dimension. At endline we find evidence that mothers are 4pp more likely to be thin (namely with a BMI below 18.5) or severely malnourished, but these findings are not robust to accounting for multiple hypothesis testing.

[55] We adapted the questionnaires to our context, translated it into Hausa, and extensively piloted it to further refine its design. Our study is among the first to evaluate the impacts of a cash transfer intervention on child cognitive and motor development in Sub-saharan Africa. Only a few randomized control trials in low-income settings have measured such outcomes, and given the wide range of instruments and scales used, our results are only partially comparable to existing work. Subject to this caveat, we note that our documented impact on communication skills is similar in size to what has been found for domains such as vocabulary and memory [Paxson and Schady 2010, Macours *et al.* 2012, Levere *et al.* 2016], and our null finding on gross motor skills is consistent with Macours *et al.* [2012] and Levere *et al.* [2016].

[56] We also examined impacts by gender of the new child. The impacts on communication skills are present for both. There is no impact for either gender on motor or personal-social skills.

A.3 Within Household Spillovers

We exploit our old and end child surveys to measure spillovers onto older and younger siblings of the new child. As Table 2 showed, the old child is on average age three at baseline, when the household begins to receive information and cash transfers. They thus lie outside the 1000-day window usually considered for early life interventions [Victora *et al.* 2010]. For the end child (the younger sibling of the new child), the household is no longer eligible for cash transfers, but their health can be impacted through persistent knowledge gains, and the sustained net resource flows into the household documented above.[57]

In Table A14 we show ITT impacts: (i) at midline on the new child (as a benchmark); (ii) at midline for the old child (that is on average five years old by midline); (iii) at endline for the end child (that is on average 12 months old). The results on birth spacing in Table 4 implied that in treated villages, the end child was born just under a month earlier than in Control villages. Hence in the end child specifications for anthropometrics, we control for their age non-parametrically using dummies for age ranges.

In Panel A we see that for the old child there is a reduction in the mean HAZ: this does not come from the left tail of the distribution – the likelihood to be stunted or severely stunted does not change significantly. Rather there might be some reallocation of resources away from old children that would have been taller, all else equal. However, Panel B shows old children have significant improvements in their health (their health outcome index increases by $.12\sigma$ while the ITT impact on the new child was $.21\sigma$). Panels C and D confirm that some of the same channels are at play as for the new child: we see significant improvements for the old child in terms of the health behaviors of mothers towards them, and the diversity of the diet they are fed.

For the end child a similar pattern of findings emerge: there are few impacts on their height (either at the mean or at thresholds for stunting), but there are marked improvements in their health: the magnitude of this is similar to that for the new child: $.19\sigma$. On mechanisms, Panels C and D show that: (i) mothers antenatal practices index for the end child rises; (ii) the health behavior index rises (and the magnitude is slightly larger than for the new child); (iii) the diversity of the diet they are fed increases significantly.

A puzzle remains in why the HAZ of the end child shows no improvement (beyond the fact that these children are still too young for such impacts to be detected). The answer might lie in the antenatal practices index: while this improves for the end child, the magnitude of the effect is $.22\sigma$, well below the $.88\sigma$ impact documented for such practices towards the new child. There might be specific practices that are especially critical for driving physical growth in early life.

[57]Most studies argue there is a far reduced scope for catching up nutritional deficiencies outside the 1000-day window [Victora *et al.* 2010], yet work based on longitudinal data suggests such catch up is possible on both physical and cognitive dimensions of human capital [Crookston *et al.* 2010]. Adhvaryu *et al.* [2019] combine early life shocks (rainfall in the year of birth) and later random assignment to receive cash transfers in the context of Progresa, to show that such transfers can reduce the negative consequences of early life shocks on schooling and employment outcomes, and that indeed those most disadvantaged at birth gain the most from later investments.

Table A15 examines this in more detail and shows ITT impacts on: (i) specific antenatal practices of mothers towards the new child and the end child (Panel A); (ii) health behaviors of mothers toward the new, old and end child. On antenatal practices we see that while the end child is more likely than the new child to be put to breast immediately, they are less likely to be born at a health facility (11pp vs. 26pp), and less likely to be exclusively breast-fed for the first six months of life (6pp vs. 29pp). These channels offer new insights on specific pathways through which antenatal practices can translate into physical development in the first 12-months of life.

On health outcomes (for which we saw more uniform impacts on the new, old and end child), Panel B of Table A15 reassuringly shows similar health behaviors towards these three groups of children in terms of being given oral rehydration salts if they have diarrhea, and given deworming medications in the six months prior to survey date.

References

[1] ADHVARYU.A, T.MOLINA, A.NYSHADHAM AND J.TAMAYOM (2019) Helping Children Catch Up: Early Life Shocks and the Progresa Experiment, mimeo, Michigan.

[2] AHMED.A, J.HODDINOTT AND S.ROY (2019) Food Transfers, Cash Transfers, Behavior Change Communication and Child Nutrition: Evidence from Bangladesh, IFPRI DP01868.

[3] ALDERMAN.H, J.R.BEHRMAN AND C.PUETT (2016) "Big Numbers about Small Children: Estimating the Economic Benefits of Addressing Undernutrition," *World Bank Research Observer* 32: 107-25.

[4] ALMAS.I, J.HAUSHOFER AND J.SHAPIRO (2019) The Income Elasticity for Nutrition: Evidence from Unconditional Cash Transfers in Kenya, NBER WP25711.

[5] ALMOND.D AND J.CURRIE (2011) "Human Capital Development Before Age Five," in O.Ashenfelter and D.Card (eds.) *Handbook of Labor Economics* Vol.4b, Elsevier.

[6] ANDERSON.M.L (2008) "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association* 103: 1481-95.

[7] ANGELUCCI.M, G.DE GIORGI AND I.RASUL (2018) "Consumption and Investment in Resource Pooling Family Networks," *Economic Journal* 128: 2613-51.

[8] ATTANASIO.O.P, V.OPPEDISANO AND M.VERA-HERNÁNDEZ (2015) "Should Cash Transfers Be Conditional? Conditionality, Preventive Care, and Health Outcomes," *American Economic Journal: Applied Economics* 7: 35-52.

[9] ATTANASIO.O.P, S.CATTAN, E.FITZSIMONS, C.MEGHIR AND M.RUBIO-CODINA (2019) "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia," *American Economic Review*, forthcoming.

[10] BAIRD.S, J.H.HICKS, M.KREMER AND E.MIGUEL (2016) "Worms at Work: Long-run Impacts of a Child Health Investment," *Quarterly Journal of Economics* 131: 1637-80.

[11] BAIRD.S, B.OZLER AND C.MCINTOSH (2019) "When the Money Runs Out: Do Cash Transfers Have Sustained Effects on Human Capital Accumulation?" *Journal of Development Economics* 140: 169-85.

[12] BANDIERA.O, R.BURGESS, N.DAS, S.GULESCI, I.RASUL AND M.SULAIMAN (2017) "Labor Markets and Poverty in Village Economies," *Quarterly Journal of Economics* 132: 811-70.

[13] BANERJEE.A.V, E.DUFLO, N.GOLDBERG, D.KARLAN, R.OSEI, W.PARIENTE, J.SHAPIRO, B.THUYSBAERT AND C.UDRY (2015) "A Multi-faceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science* 348: Issue 6236.

[14] 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: 86-125.

[15] BERGMAN.P, R.CHETTY, S.DELUCA, N.HENDREN, L.F.KATZ AND C.PALMER (2019) Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice, NBER WP 26164.

[16] BERNHARDT.A, E.FIELD, R.PANDE AND N.RIGOL (2019) "Household Matters: Revisiting the Returns to Capital among Female Microentrepreneurs," *AER: Insights* 1: 141-60.

[17] BHUTTA.Z ET AL. (2008) "What Works? Interventions for Maternal and Child Undernutrition and Survival," *Lancet* 371: P417-40.

[18] BHUTTA.Z ET AL. (2013) "Evidence Based Interventions for Improvement of Maternal and Child Nutrition: What can be Done and at what Cost?," *Lancet* 382: 452-77.

[19] BOUGUEN.A, Y.HUANG, M.KREMER AND E.MIGUEL (2019) "Using RCT's to Estimate Long-Run Impacts in Development Economics," *Annual Review of Economics*, forthcoming.

[20] BROWNING.M, P.CHIAPPORI AND V.LECHENE (2010) "Distributional Effects in Household Models: Separate Spheres and Income Pooling," *Economic Journal* 120: 786-99.

[21] CAMPBELL.F, G.CONTI, J.J.HECKMAN, S.H.MOON, R.PINTO, E.PUNGELLO AND Y.PAN (2014) "Early Childhood Investments Substantially Boost Adult Health," *Science* 343: 1478-85.

[22] CAEYERS.B, S.KRUTIKOVA AND O.ATTANASIO (2016) +Parent Support in the Income Sector: A Review of the Evidence, IFS Report.

[23] CONTI.G, J.J.HECKMAN AND R.PINTO (2016) "The Effects of Two Influential Early Childhood Interventions on Health and Healthy Behaviour," *Economic Journal* 126: F28-65.

[24] CROOKSTON.B.T ET AL. (2010) "Children who Recover from Early Stunting and Children who are not Stunted Demonstrate Similar Levels of Cognition," *Journal of Nutrition* 140: 1996-2001.

[25] CURRIE.J AND M.ROSSIN-SLATER (2013) "Weathering the Storm: Hurricanes and Birth Outcomes," *Journal of Health Economics* 32: 487-503.

[26] DASGUPTA.P (1997) "Nutritional Status, the Capacity for Work, and Poverty Traps," *Journal of Econometrics* 77: 5-37.

[27] DEAN.E.B, H.SCHOFIELD AND F.SCHILBACH (2018) "Poverty and Cognitive Function," in C.Barrett, M.Carter and J-Paul.Chavas (eds.) *The Economics of Asset Accumulation and Poverty Traps*, University of Chicago Press.

[28] DEWEY.K.G AND S.ADU-AFARWUAH (2008) "Systematic Review of the Efficacy and Effectiveness of Complementary Feeding Interventions in Developing Countries," *Maternal and Child Nutrition* 4: 24-85.

[29] DHALIWAL.I, E.DUFLO, R.GLENNERSTER AND C.TULLOCH (2012) A Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education, mimeo, MIT.

[30] DORSEY.E.R, E.OSTER AND I.SHOULSON (2013) "Optimal Expectations and Limited Medical Testing: Evidence from Huntington's Disease," *American Economic Review* 103: 804-30.

[31] DOYLE.O (2019) "The First 2000 Days and Children's Skills," *Journal of Political Economy*, forthcoming.

[32] ELSHIBLY.E.M AND G.SCHMALISCH (2008) "The Effect of Maternal Anthropometric Characteristics and Social Factors on Gestational Age and Birth Weight in Sudanese Newborn Infants," *BMC Public Health* 8: 1-7.

[33] EVANS.W.N AND D.S.LIEN (2005) "The Benefits of Prenatal Care: Evidence from the PAT Bus Strike," *Journal of Econometrics* 125: 207-39.

[34] FERNALD.L.C.A, R.M.C.KAGAWA, H.A.KNAUER AND L.SCHNASS (2017) "Promoting Child Development Through Group-Based Parent Support Within a Cash Transfer Program: Experimental Effects on Children's Outcomes," *Developmental Psychology* 53: 222-36.

[35] FINK.G ET AL. (2016) "Schooling and Wage Income Losses Due to Early-Childhood Growth Faltering in Developing Countries: National, Regional, and Global Estimates," *The American Journal of Clinical Nutrition* 104: 104-12.

[36] GELBACH.J.B (2016) "When do Covariates Matter? And Which Ones, and how Much?," *Journal of Labor Economics* 34: 509-43.

[37] GERTLER.P, J.J.HECKMAN, R.PINTO, A.ZANOLINI, C.VERMEERSCH, S.WALKER, S.M.CHANG AND S.GRANTHAM-MCGREGOR (2014) "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica," *Science* 344: 998-1001.

[38] GRANTHAM-MCGREGOR.S ET AL. (2007) "Developmental Potential in the First 5 Years for Children in Developing Countries," *Lancet* 369: 60-70.

[39] HANDA.S, D.SEIDENFELD, B.DAVIS, G.TEMBO AND ZAMBIA CASH TRANSFER EVALUATION TEAM (2016) "The Social and Productive Impacts of Zambia's Child Grant: Impacts of Zambia's Child Grant," *Journal of Policy Analysis and Management* 35: 357-87.

[40] HEADEY.D.D, K.HIRVONEN AND J.F.HODDINOTT (2018) "Animal Sourced Foods and Child Stunting," *American Journal of Agricultural Economics* 100: 1302-19.

[41] HECKMAN.J.J, S.MOON, R.PINTO, P.SAVELYEV AND A.YAVITZ (2010) "The Rate of the Return to the HighScope Perry Preschool Program," *Journal of Public Economics* 94: 114-28.

[42] HECKMAN.J.J AND G.KARAPAKULA (2019) Intergenerational and Intragenerational Externalities of the Perry Preschool Project, NBER WP25889.

[43] HODDINOTT.J, J.A.MALUCCIO, J.R.BEHRMAN, R.FLORES AND R.MARTORELL (2008) "Effect of a Nutrition Intervention During Early Childhood on Economic Productivity in Guatemalan Adults," *The Lancet* 371: 411-16.

[44] HODDINOTT.J ET AL. (2011) The Consequences of Early Childhood Growth Failure over the Life Course, IFPRI DP 01073.

[45] HODDINOTT.J ET AL. (2013) "Adult Consequences of Growth Failure in Early Childhood," *American Journal of Clinical Nutrition* 98: 1170-8.

[46] IMAI.K, L.KEELE AND D.TINGLEY (2010) "A General Approach to Causal Mediation Analysis," *Psychological Methods* 15: 309-34.

[47] KAKIETEK.J, J.D.EBERWEIN, D.WALTERS AND M.SHEKAR (2017) Unleashing Gains in Economic Productivity with Investments in Nutrition, Washington, DC: World Bank Group.

[48] KRAMER.M.S AND R.KAKUMA (2012) "Optimal Duration of Exclusive Breastfeeding," *Cochrane Database of Systematic Reviews* 8.

[49] KREMER.M AND R.GLENNERSTER (2012) "Improving Health in Developing Countries: Evidence from Randomized Evalautions," in M.V.Pauly, T.G.Mcguire and P.Barros (eds.) *Handbook of Health Economics Vol.2*, New York: Elsevier.

[50] LEVERE.M, G.ACHARYA, P.BHARADWAJ (2016) The Role of Information and Cash Transfers on Early Childhood Development: Evidence from Nepal, NBER WP 22640.

[51] LI.K AND D.J.POIRIER (2003) "Bayesian Analysis of an Econometric Model of Birth Inputs and Outputs," *Journal of Population Economics* 16: 597-625.

[52] MACOURS.K, N.SCHADY AND R.VAKIS (2012) "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment," *American Economic Journal: Applied Economics* 4: 247-73.

[53] MALUCCIO.J.A AND R.FLORES (2004) Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social, FCND DP184.

[54] MANI.A, S.MULLAINATHAN, E.SHAFIR AND J.ZHAO (2013) "Poverty Impedes Cognitive Function," *Science* 341: 976-80.

[55] MANLEY.J, S.GITTER AND V.SLAVCHEVSKA (2013) "How Effective are Cash Transfers at Improving Nutritional Status?," *World Development* 48: 133-55.

[56] OKEKE.T.C, J.O.ENWEREJI, H.U.EZEGWUI, A.C.ADIUKU-BROWN, C.O.ADIRI AND O.A.OKEZIE (2014) "Prospective Estimation of Duration of Pregnancy in Women Attending Antenatal Care at the University of Nigeria Teaching Hospital, Enugu, Nigeria," *Annual Research and Review in Biology* 4: 2274-82.

[57] PAXSON.C AND N.SCHADY (2010) "Does Money Matter? The Effects of Cash Transfers on Child Development in Rural Ecuador," *Economic Dev. and Cultural Change* 59: 187-229.

[58] RICHTER.L ET AL. (2017) "Investing in the Foundation of Sustainable Development: Pathways to Scale Up for Early Childhood Development," *Lancet* 389: 103-18.

[59] ROMANO.J.P AND M.WOLF (2005) "Stepwise Multiple Testing as Formalized Data Snooping," *Econometrica* 73: 1237-82.

[60] SHARP.K, A.CORNELIUS AND V.GADHAVI (2018) *Child Development Grant Programme Evaluation Qualitative Endline Report*, Technical report, ePact Consortium.

[61] SQUIRES.J AND D.D.BRICKER (2009) *Ages & Stages Questionnaires*, Brookes Publishing Company OCLC: 1010500088.

[62] SRIDHAR.D AND A.DUFFIELD (2006) A Reiew of the Impact of Cash Transfer Programmes on Child Nutritional Status and Some Implications for Save the Children UK Programmes, Save the Children Report.

[63] THOMAS.D AND J.STRAUSS (1997) "Health and Wages: Evidence on Men and Women in Urban Brazil," *Journal of Econometrics* 77: 159-85.

[64] VICTORA.C.G, M.DE ONIS, P.C.HALLAL, M.BLOSSNER AND R.SHRIMPTON (2010) "Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions," *Pediatrics* 125: e473-80.

[65] WHO (1995) *Physical Status: Uses and Interpretation of Anthropometry*, WHO Expert Committee on Nutrition, Geneva.

[66] WHO (2008) *Indicators for Assessing Infant and Young Child Feeding Practices*, Washington DC.: WHO, OCLC: 795844943.

[67] WHO (2009) *WHO Child Growth Standards: Growth Velocity Based on Weight, Length and Head Circumference: Methods and Development*, Geneva, Swtizerland: WHO Department of Nutrition for Health and Development, OCLC: ocn316014678.

Table 1: Baseline Balance - Village, Household and Parental Characteristics

Sample: Households with pregnant women at baseline (N=3688)
Means, standard deviation in braces, p-values in brackets

	(1) Control	(2) T1	(3) T2	(1) = (2)	(1) = (3)	(2) = (3)
Panel A: Village						
Observations	68	71	71			
Experienced natural shock in past year	82.4	84.5	84.5	[.706]	[.823]	[.585]
Experienced man made shock in past year	35.3	38.0	43.7	[.629]	[.230]	[.566]
Distance to closest market (km)	1.86	2.68	1.86	[.141]	[.943]	[.096]
	{2.38}	{2.49}	{2.17}			
Distance to closest health facility (km)	1.44	1.49	1.58	[.719]	[.567]	[.949]
	{1.53}	{1.44}	{1.61}			
Panel B: Household						
Observations	1186	1258	1244			
Household size	7.70	7.43	7.55	[.406]	[.492]	[.905]
	{4.33}	{4.27}	{4.37}			
Number of children aged 0-18	4.63	4.46	4.59	[.562]	[.837]	[.729]
	(3.25)	{3.24}	{3.36}			
Monthly food expenditure (in $USD)	84.9	90.6	79.9	[.158]	[.537]	[.041]
	(122)	(129)	(118)			
Share of monthly expenditures on food (%)	38.5	38.4	38.4	[.709]	[.588]	[.819]
	{.255}	{.257}	{.256}			
Living on less than $1.90/ day (extreme poverty) (%)	72.2	69.8	73.6	[.145]	[.444]	[.029]
Total monthly earnings	189	203	184	[.244]	[.86]	[.262]
	{381}	{371}	{378}			
Did not have enough food in past year (%)	15.7	13.4	16.2	[.491]	[.781]	[.317]
Household owns any animals (%)	73.3	69.6	70.4	[.12]	[.239]	[.69]
Panel C: Women						
Observations	1186	1258	1244			
Age (years)	25.5	25.3	25.2	[.849]	[.324]	[.469]
	{6.82}	{6.81}	{6.90}			
Can read and write at least one language (%)	19.1	20.5	22.1	[.383]	[.398]	[.766]
Polygamous relationship (%)	49.1	46.3	51.1	[.271]	[.428]	[.058]
Paid/unpaid work in past year (%)	74.3	69.5	70.5	[.491]	[.358]	[.822]
Total monthly earnings (in $USD)	25.7	25.0	24.0	[.92]	[.419]	[.589]
	{49.9}	{44.2}	{45.5}			
Rearing/ tending or selling household livestock (%)	36.8	33.7	28.7	[.829]	[.06]	[.053]
Panel D: Husband						
Observations	952	916	912			
Age (years)	43.0	42.4	42.1	[.415]	[.076]	[.388]
	{9.12}	{9.42}	{9.28}			
Can read and write in at least one language (%)	42.9	68.8	38.9	[.206]	[.846]	[.060]
Paid/unpaid work in past year (%)	94.0	95.0	92.7	[.234]	[.500]	[.041]
Total monthly earnings (in $USD)	163	177	160	[.222]	[.969]	[.298]
	{371}	{361}	{368}			
Farming household's land (%)	81.9	81.3	78.8	[.982]	[.366]	[.431]

Notes: Panel A reports data from the village surveys, and Panels B, C and D report data from the household surveys. In Panel A, Natural shock in the village in the past year is a dummy equal to one if the village experiences a drought, flood, crop damage by pests or by disease. A Man made shock in the village in the past year is a dummy equal to one if the village experiences curfews, land disputes, violence, widespread migration or cattle rustling. Distance measures are the straight line distances in KM. In Panel B, household size is the number of people living in the household with common eating arrangements. Food expenditure is based on 7-day recall for food items. Total expenditure is based on: food expenditure, a 7-day recall for consumable items (e.g. petrol, fuel, phone credit, cigarettes), a 30-day recall for items such as toiletries and clothing and an annual recall for larger items such as dowry, funerals and school expenses as well as durables such as mattress, table motorbike, which we then convert to a monthly expenditure measure. Living on less than $1.90 a day indicates if the household is spending less than $1.90 a day according to PPP USD in 2011 terms. This is the World Bank's international poverty line definition for households residing in extreme poverty. In Panels C and D, Total monthly earnings are the earnings for the husband and wife reported from the past year across all work activities that are carried out for pay. Values above the 99th percentile are set to missing. Columns 1 to 3 report the mean (and standard deviation for continuous variables) of the variable in the Control group, the low-intensity information treatment arm (T1) and the high-intensity information treatment arm (T2). The p-values on tests of equality across Columns are obtained from an OLS regression, controlling for randomization stratum and clustering standard errors at the village level. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table 2: Baseline Balance - Child Related Outcomes

Sample: Households with pregnant women at baseline (N=3688)
Means, standard deviation in braces, p-values in brackets

	(1) Control	(2) T1	(3) T2	(1) = (2)	(1) = (3)	(2) = (3)
Panel A: Wife's Knowledge						
Health facility is best place to give birth (%)	15.6	16.2	14.3	[.734]	[.397]	[.669]
Should breastfeed exclusively for 6 months (%)	13.5	16.9	15.2	[.225]	[.603]	[.359]
Panel B: Husband's Knowledge						
Health facility is best place to give birth (%)	20.7	20.4	19.4	[.592]	[.388]	[.713]
Should breastfeed exclusively for 6 months (%)	12.8	14.2	12.4	[.795]	[.553]	[.316]
Panel C: Old Child						
Observations	844	877	876			
Age (Months)	36.4	36.8	36.3	[.48]	[.925]	[.372]
	{11.5}	{11.7}	{11.7}			
Birth spacing from previous child	27.4	27.7	26.9	[.719]	[.354]	[.184]
	{10.3}	{10.8}	{11.0}			
Child put to the breast immediately (%)	31.6	28.6	30.1	[.474]	[.752]	[.526]
Appropriately breastfed (%)	25.3	28.6	18.9	[.940]	[.376]	[.341]
Had diarrhea in past 2 weeks (%)	32.7	26.4	28.4	[.001]	[.044]	[.167]
If had diarrhea in past two weeks:						
Anyone sought advice/treatment (%)	79.5	79.5	78.4	[.869]	[.623]	[.485]
Given ORS for diarrhea (%)	38.3	46.4	38.0	[.078]	[.723]	[.081]
Stunted (HAZ<-2) (%)	69.0	65.9	68.8	[.135]	[.801]	[.210]
Wasted (WHZ<-2) (%)	6.1	5.8	6.0	[.891]	[.755]	[.958]
Panel D: New Child (in utero at Baseline)						
Observations	1670	1,736	1,750			
Month of pregnancy	5.27	5.315	5.203	[.530]	[.362]	[.143]
	{2.18}	{2.15}	{2.14}			

Notes: In Panel C, the Old Child variable child put to breast immediately is a dummy for the child having been put to the breast in the first 30 minutes after birth. The appropriately breastfed variable is a dummy indicating age-appropriate breastfeeding according to WHO guidelines [WHO 2008, i.e. exclusive breastfeeding up to the age of 6 months and complementary breastfeeding from 6 to 23 months. Stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -2 standard deviations of the WHO defined guidelines [WHO 2009]. Wasted is a dummy indicating children with weight-for-height-z-score (WHZ) under -2 standard deviations of the WHO defined guidelines. In Panel D, the New Child month of pregnancy variable is reported by mothers pregnant at Baseline. Columns 1 to 3 report the mean (and standard deviation for continuous variables) of the variable in the Control group, the low-intensity information treatment arm (T1) and the high-intensity information treatment arm (T2). The p-values on tests of equality across Columns are obtained from an OLS regression, controlling for randomization stratum and clustering standard errors at the village level.

Table 3: Take-up of Cash Transfers

Sample: Households with pregnant women at baseline (N=3688)
Means, standard deviation in parentheses, p-values in brackets

	(1) Control	(2) T1	(3) T2	(2) = (3)
Panel A: Receipt				
Ever received transfer	10.8	90.3	92.8	[.485]
Panel B: Timing of First Transfer				
Age of new child (in utero) at first payment (months)		-1.05	-.538	[.695]
		{9.34}	{9.37}	
During pregnancy (%)		41.9	44.5	[.311]
1st trimester (%)		3.5	3.0	[.734]
2nd trimester (%)		13.1	11.2	[.305]
3rd trimester (%)		25.4	30.3	[.046]
In month of birth (%)		13.8	11.5	[.183]
After birth (%)		33.1	33.7	[.932]
Panel C: Intensity of Treatment				
Number of payments		24.0	23.3	[.192]
		{5.69}	{5.74}	
Total amount transferred		477	464	[.206]
		{121}	{124}	
Receiving payments at midline (%)		83.6	83.4	[.872]
Receiving payments at endline (%)		8.80	10.00	[.826]

Notes: This uses data from the administrative records data on payments. The age of the new child at first payment is derived from the month of pregnancy as reported by mothers pregnant at Baseline. Columns 1 to 3 report the mean (and standard deviation for continuous variables) of the variable in the Control group, the low-intensity information treatment arm (T1) and the high-intensity information treatment arm (T2). The p-values on tests of equality across Columns are obtained from an OLS regression, controlling for randomization stratum and clustering standard errors at the village level. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table 4: Fertility, Birth Spacing and Child Mortality

Sample: Households with pregnant women at baseline (N=3688)

Column 1: Standard deviation in braces

Columns 2 and 3: Standard errors in parentheses clustered by village

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline
Panel A: Fertility			
Any child born (%)	84.8	1.92	-2.62
		(1.34)	(1.75)
Number of children born	1.06	.022	.022
	{.433}	(.015)	(.022)
Panel B: Gestation and Birth Spacing			
Month of birth of new child		.524*	
		(.302)	
Birth spacing between new and Endline child (months)	31.5		-.806**
	{5.81}		(.339)
Birth spacing between new and Endline child <= 24 months (%)	7.60		3.27**
			(1.54)
Panel C: Mortality			
Any child born between Baseline and Midline that died (%)	12.9	.650	-.598
		(1.24)	(1.09)
Age of death (months)	6.50	.916	-.643
	{6.84}	(.712)	(.773)
Died age < 6 months (%)	.486	-.066	.050
		(.051)	(.067)

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) value in Control households, at Midline, based on fertility since Baseline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure and a dummy for polygamous relationships. In Panel B we also control for the month in utero of the new child. Standard errors are clustered at the village level throughout. Any child born refers to a child born between Baseline and Midline (Midline and Endline) in Column 2 (3). The age child death outcomes relate to children born between Baseline and Midline.

Table 5: Child Outcomes

Sample: Households with pregnant women at baseline (N=3688)
Columns 1 and 2: Standard deviation in braces
Columns 2 and 4: Standard errors in parentheses clustered by village
Romano and Wolf [2016] Bootstrap p-values in square brackets below standard errors

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Panel A: Height and Stunting				
Height-for-Age (HAZ)	-2.46	.198***	.119**	[.236]
	{1.30}	(.068)	(.059)	
		[.010]	[.103]	
Stunted (HAZ < -2) (%)	66.2	-5.22**	-4.87*	[.897]
		(2.43)	(2.55)	
		[.011]	[.103]	
Severely stunted (HAZ < -3) (%)	34.8	-4.77**	-4.21**	[.812]
		(2.21)	(2.15)	
		[.103]	[.103]	
Panel B: Health Outcomes				
Health Outcome Index	.000	.209***	.288***	[.204]
	{1.00}	(.052)	(.050)	
Been ill/injured in last month (%)	69.6	-8.53***	-12.0***	[.269]
		(2.36)	(2.39)	
Had diarrhea in past two weeks (%)	37.8	-6.90***	-9.30***	[.417]
		(2.21)	(2.36)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) value in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships, and the gender of the new child. Standard errors are clustered at the village level throughout. We also report p-values adjusted for multiple testing in square brackets below the standard errors. These are computed using the step-down procedure discussed in Romano and Wolf [2016], with 1,000 bootstrap replications. There are three outcomes being simultaneously tested in columns 2 and 3 – for each outcome we test the impact at midline and endline. Therefore, the p-values in columns 2 to 3 are adjusted for testing on 6 hypotheses. Stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -2 standard deviations of the WHO defined guidelines [WHO 2009]. Severely stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -3 standard deviations of the WHO defined guidelines. The Health Outcome Index is constructed as in Anderson [2008], and standardized to have mean zero and variance one in the Control group at Midline. The index includes the following health outcome components: a dummy variable that takes the value of 1 if the child has not been ill in the last month and a dummy variable that takes the value of 1 if the child has not had diarrhoea in the past two weeks.

Table 6: Parental Knowledge

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village
Romano and Wolf [2016] Bootstrap p-values in square brackets below standard errors

	Wife				Husband				Wife = Husband	
	(1) Control Mean	**(2) ITT, Midline**	**(3) ITT, Endline**	**(2) = (3)**	**(4) Control Mean**	**(5) ITT, Midline**	**(6) ITT, Endline**	**(5) = (6)**	**(2) = (5)**	**(3) = (6)**
Knowledge index	.000	.951***	.796***	[.027]	.000	.382***	.256***	[.023]	[.000]	[.000]
	{1.00}	(.088)	(.089)		{1.00}	(.045)	(.045)			
		[.001]	[.001]			[.001]	[.001]			
Panel A: Prenatal										
Would advise to seek a check-up, even if healthy (%)	66.8	7.83***	5.40***	[.189]	71.8	4.65***	2.00	[.200]	[.104]	[.240]
		(1.99)	(1.22)			(1.79)	(1.39)			
Panel B: Perinatal										
Colostrum is good for the baby (%)	63.4	19.3***	15.6***	[.087]	67.2	16.1***	13.9**	[.548]	[.258]	[.527]
		(2.63)	(2.44)			(4.61)	(5.54)			
Best to start breastfeeding immediately (%)	18.2	26.7***	18.6***	[.012]	20.0	12.3***	12.3***	[.985]	[.001]	[.083]
		(2.79)	(2.48)			(2.99)	(3.74)			
Best place to give birth is health facility (%)	13.5	28.7***	26.5***	[.430]	20.7	11.0***	17.6***	[.081]	[.680]	[.791]
		(3.58)	(3.85)			(3.24)	(3.35)			
Baby should not receive other liquids on first day (%)	46.2	21.7***	25.4***	[.250]	49.5	19.2***	18.0***	[.785]	[.278]	[.097]
		(2.80)	(3.03)			(3.34)	(3.30)			
Panel C: Postnatal										
Do not give baby water when hot outside (%)	8.60	39.2***	42.3***	[.284]	9.60	25.5***	31.9***	[.161]	[.000]	[.013]
		(3.38)	(3.38)			(2.86)	(3.68)			
Never give water to a baby under 6 months old (%)			44.8***				26.8***			[.000]
			(3.21)				(3.24)			
Best to breastfeed exclusively for 6 months (%)	13.5	28.8***	26.5***	[.430]	12.8	11.9***	6.85***	[.014]	[.000]	[.000]
		(3.58)	(3.84)			(1.84)	(1.65)			

Notes: Significance levels: * (10%), ** (5%), ***(1%). Columns 1 and 4 show the mean (and standard deviation for continuous outcomes) values in Control households at Baseline. Columns 2 and 5 report ITT estimates at Midline, and Columns 3 and 6 report ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships, and the gender of the new child. Standard errors are clustered at the village level throughout. We also report p-values adjusted for multiple testing in square brackets below the standard errors for certain outcomes. These are computed using the step-down procedure discussed in Romano and Wolf [2016], with 1,000 bootstrap replications. There is one outcome being simultaneously tested at midline and endline for wives and husbands. Therefore, the p-values are adjusted for testing on 4 hypotheses. In the first row of the table, the Knowledge indices are constructed as in Anderson [2008], and standardized to have mean zero and variance one in the Control group at Baseline. Each index includes the following question components: Would you advise to seek a check-up even if the baby is healthy? Is colostrum good for the baby? Should you breastfeed immediately? Where is best place to give birth? Should a baby receive any other liquids on first day? Should you give water to a baby if it is hot out? How long should you exclusively breastfeed for?

Table 7: Mother's Practices and Health Behaviors

Sample: Households with pregnant women at baseline (N=3688)

Standard deviation in braces

Standard errors in parentheses clustered by village

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Practices index	.000	.841***		
	{1.00}	(.085)		
Panel A: Prenatal				
Had antenatal care (%)	19.5	16.8***		
		(3.26)		
Panel B: Perinatal				
Fed colostrum in first hour (%)	38.1	29.2***		
		(2.96)		
Born at health facility (%)	13.0	4.75**		
		(2.00)		
Put to breast immediately (%)	44.3	26.3***		
		(3.14)		
Panel C: Postnatal				
Exclusively breastfed for 6 months (%)	11.7	29.0***		
		(2.87)		
Panel D: Health Behaviors				
Health Behavior Index	.000	.173***	.333***	[.082]
	{1.00}	(.057)	(.068)	
Given deworming medication in past 6 months (%)	16.4	8.02***	12.1***	[.312]
		(2.43)	(2.92)	
Received all basic vaccinations (%)	.809	.541	2.83***	[.035]
		(.553)	(.971)	
If had diarrhea in past two weeks:				
Anyone sought advice/treatment (%)	78.3	6.88**	7.57**	[.887]
		(3.02)	(3.52)	
Given ORS for diarrhea (%)	40.8	10.3**	14.1***	[.532]
		(4.03)	(4.77)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout. In the first row of the table, the Practices index is constructed as in Anderson [2008], and standardized to have mean zero and variance one in the Control group at Baseline. The index includes the following question components: Did the child receive antenetal care? Was the child fed colostrum in the first hour? Was the child put to breast immediately? Was the child born at a health facility? And (if applicable) was the child exclusively breastfed for 6 months? In Panel D, the Health behaviour index is similarly constructed from the following question components: Has the child been given deworming medicine in the past 6 months? Has the child received all basic vaccinations? The received all basic vaccinations outcome is a dummy equal to one if the child has received the following vaccinations: BCG, three polio vaccinations, three DPT vaccinations and measles. Vaccinations are acknowledged from the child having a vaccination card (or it being reported on their birth card).

Table 8: Dietary Diversity and Food Security

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village
Romano and Wolf [2016] Bootstrap p-values in square brackets below standard errors

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Panel A: Dietary Diversity				
Dietary diversity index	3.22	.364***	.373***	[.915]
	{1.49}	(.071)	(.071)	
		[.001]	[.001]	
4+ MDD groups (%)	46.6	10.7***	12.7***	[.536]
		(2.45)	(2.44)	
		[.002]	[.001]	
Panel B: Food Security (in past 30 days)				
Did not have enough food (%)	16.4	-4.39***	-9.34***	[.020]
		(1.60)	(1.91)	
		[.125]	[.001]	
Ever went whole day and night without eating (%)	4.94	-.932	-4.76***	[.014]
		(.822)	(1.41)	
		[.490]	[.002]	
Ever went to bed hungry (%)	8.80	-2.37**	-6.08***	[.011]
		(1.17)	(1.35)	
		[.125]	[.002]	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. In addition, we control for the gender of the child and the child's age in months. Standard errors are clustered at the village level throughout. We also report p-values adjusted for multiple testing in square brackets below the standard errors for certain outcomes. These are computed using the step-down procedure discussed in Romano and Wolf [2016], with 1,000 bootstrap replications. There are two outcomes in Panel A being simultaneously tested at midline and endline. Therefore, the p-values are adjusted for testing on 4 hypotheses. In Panel B there are three outcomes being simultaneously tested at midline and endline. Therefore, the p-values are adjusted for testing on 6 hypotheses. The diet diversity index and the food groups are obtained from a 24-hour food recall module administered to the child's mother or main career. Each meal consumed in the day before the interview from waking up to bedtime is recorded, and each ingredient is coded into categories. The Dietary Diversity Index sums the number of food groups the child has received from the following 7 food groups: 1. Grains, roots and tubers, 2. Legumes and nuts, 3. Dairy products, 4. Flesh foods, 5. Eggs, 6. Vitamin-A rich fruits and vegetables, 7. Other fruits and vegetables. 4 + MDD groups means the child has eaten in at least 4 of the mentioned 7 food groups.

Table 9: Seasonal Food Security and Coping Strategies

Sample: Households with pregnant women at baseline (N=3688)
Standard errors in parentheses clustered by village

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Did not have enough food in past year (%)	28.6	-6.26***	-11.55***	[.044]
		(2.38)	(2.42)	
during Kaka (Mid Oct to Dec)	4.16	-2.29***	-2.14**	[.884]
		(.847)	(.833)	
during Sanyi (Dec to Feb)	5.15	-3.61***	-3.24***	[.721]
		(.969)	(.795)	
during Rani (Mar to May)	15.7	-5.99***	-5.97***	[.990]
		(1.52)	(1.53)	
during Damuna (Jun to Mid Oct)	20.1	-4.00**	-11.48***	[.003]
		(1.95)	(2.35)	
Why not enough food? (%)				
Food too expensive/didn't have enough money	21.2	-6.02***	-10.12***	[.104]
		(1.91)	(2.16)	
Unable to reach the market	10.0	-3.18**	-4.24***	[.512]
		(1.37)	(1.33)	
Small land size	8.33	-2.16	-3.82***	[.273]
		(1.34)	(1.15)	
Lack of farm inputs	5.45	-1.18	-2.15**	[.473]
		(1.09)	(.891)	
Strategy to deal with not enough food? (%)				
Helped by relatives or friends	12.3	-4.22**	-5.33***	[.557]
		(1.68)	(1.26)	
Took on more work	12.2	-4.80***	-4.56***	[.887]
		(1.52)	(1.37)	
Reduced condiments and sauces in meals	6.44	-2.86***	-3.56***	[.664]
		(1.06)	(1.12)	
Borrowed money	5.25	-1.89**	-1.60*	[.803]
		(.906)	(.898)	
Household members moved away to find work	3.57	-2.39***	-2.30***	[.926]
		(.725)	(.672)	
Sold livestock	2.48	-0.69	-2.38***	[.087]
		(.676)	(.786)	
Ate limited range of food	1.09	-.105	-2.35***	[.007]
		(.359)	(.757)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout.

Table 10: Labor Activities

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village

	Wife				Husband			
	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)	(4) Control Mean	(5) ITT, Midline	(6) ITT, Endline	(5) = (6)
Panel A: Labor Activities								
Any work in past year (%)	74.3	6.12**	11.0***	[.132]	94.5	.287	-.306	[.954]
		(2.53)	(1.96)			(.211)	(.222)	
Multiple activities (%)	30.5	2.73	12.8***	[.221]	62.5	2.49	4.72*	[.453]
		(3.28)	(3.17)			(2.23)	(2.44)	
Days/week working in highest-earning activity	2.63	.225	.575***	[.193]	3.60	-.011	.404*	[.085]
	{3.06}	(.165)	(.219)		{2.9}	(.141)	(.223)	
Panel B: Activity Type								
Has business/self-employed (%)	55.7	5.66**	11.2***	[.130]	47.7	-2.97	2.62	[.018]
		(2.70)	(2.85)			(2.31)	(1.96)	
Petty trading (%)	42.6	4.78*	10.3***	[.149]	-	-	-	
		(2.85)	(2.92)					
Farming own land (%)	-	-	-		81.5	-.651	.218	[.867]
						(.950)	(.683)	
Panel C: Investment								
Monthly expenditure on wife's business inputs	-	-	21.4***					
			(4.57)					
Monthly expenditure on husband's business inputs	-	-	-4.83					
			(4.10)					
Owning any livestock (%)	78.3	6.56***	12.1***	[.016]				
	{41.2}	(1.94)	(2.12)					
Panel D: Earnings								
Total monthly earnings from employed and self-employed activities	95.5	19.4***	19.4***	[.999]	207	13.9	8.81	[.795]
	{172}	(6.79)	(5.48)		{335}	(18.4)	(11.6)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Baseline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout. Work activities are defined as any paid or unpaid work, either self-employed or salaried, excluding housework and childcare. Self-employed activities are ones where payments are received directly from the client/customer (e.g. hairdresser working in her own shop) rather than from an employer. Panel B includes the most common labor activities that woman and husbands in our sample engage in: petty trade for women and farming their land for husbands. Panel C shows investment into the wife and husband's business inputs. Panel D shows total earnings. There are methodological differences in how earnings were measured at Midline and Endline. At Endline, we slightly changed the questionnaire to capture subtler aspects of income generating activities. For activities such as petty trading and small self-operated artisanal activities, we elicited cost of inputs and sales revenue instead of a more generic "last payment received". Total earnings are then constructed by summing payments and profits (for self-employed work). Values above the 99th percentile are set to missing. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table 11: Expenditure, Saving, Borrowing and Net Resources

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Panel A: Expenditure				
Monthly food expenditure	84.2	20.9**	16.2**	[.664]
	{120}	(9.48)	(7.05)	
Total monthly expenditure	207	40.4**	25.2*	[.437]
	{251}	-17.5	-13.3	
Share of total expenditure on food (%)	39.4	1.66	2.16*	[.736]
		(1.20)	(1.13)	
Panel B: Saving/Borrowing				
Total savings (including in kind)	272	-20.8	55.6*	[.220]
	{693}	(57.3)	(28.7)	
Total borrowed	39.5	-21.60	-19.85*	[.918]
	{169}	(14.9)	(11.6)	
Panel C: Net Resources and Extreme Poverty				
Change in monthly net resources	-	48.5**	48.9***	[.984]
		(19.4)	(18.6)	
Extreme poverty index: likelihood above extreme poverty line of $1.90/day (0-100)	26.7	.69	1.43**	[.170]
	{12.8}	(.55)	(.57)	

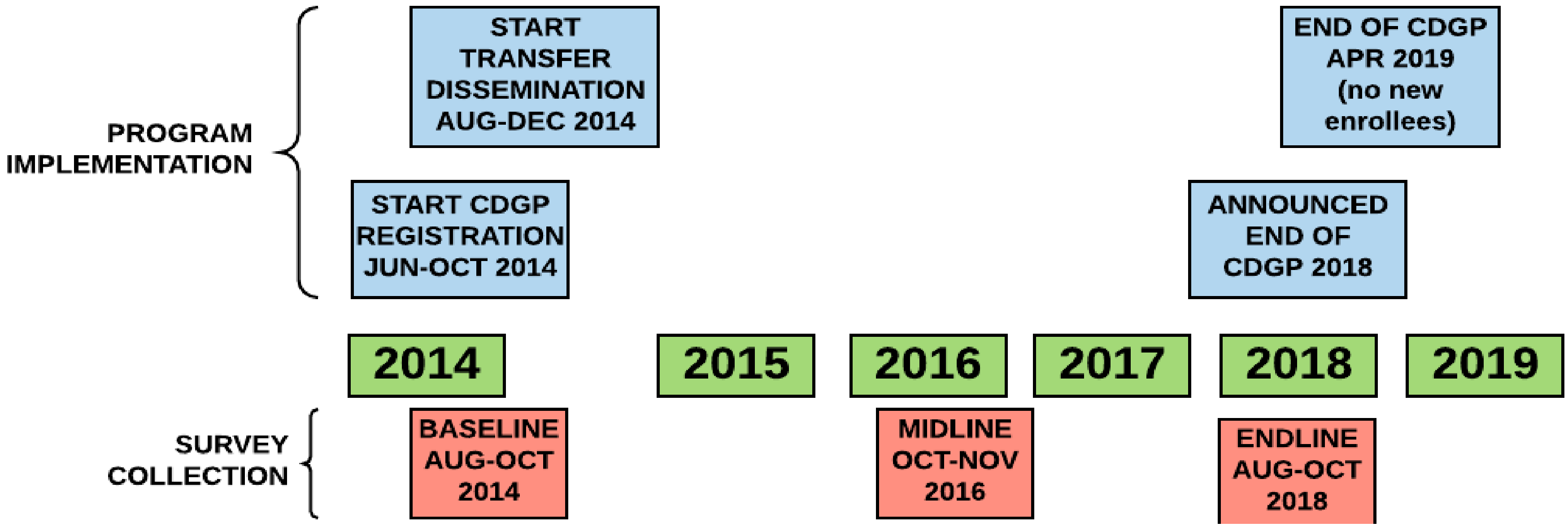
Notes: Significance levels: * (10%), ** (5%), ***(1%). In Panels A, B and C, Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Baseline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout. Food expenditure is obtained using a 7-day expenditure recall of 13 food items. Non-food expenditure is obtained combining the following sources: a 7-day expenditure recall of consumables (e.g. matches, fuel), a 30-day recall of other items (e.g. toiletries, utensils, household items, health expenditure), a 12-month recall of major expenses (e.g. school fees, ceremony costs, remittances); expenditure on durables using a 12-month recall of expenditure on assets the household owns (e.g. TV set, wheelbarrow, mattress). The top 1% of total expenditure amounts are trimmed. Net resources = income + transfers - saving + borrowing. As saving and borrowing are measured as stocks, we convert these into monthly flows assuming they accumulate at a constant rate between survey waves. The Poverty index is the Progress out of Poverty Index (PPI). For each household, the PPI is calculated through a scorecard and its value, ranging from 0 to 100, represents the likelihood a household is above the global extreme poverty line ($1.90 a day). All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table 12: Internal Rate of Return

	Cash Transfers as Purely Redistributive				Cash Transfers as Pure Cost			
	Household	**Boys**	**Girls**	**Average Child**	**Household**	**Boys**	**Girls**	**Average Child**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Social discount rate = 5%, Resource gains sustained for 5 years, earnings gains from age 16-60								
A. Cost parameters								
NPV Cash transfer					536	536	536	536
Administrative costs of cash transfers	54	54	54	54	54	54	54	54
Administrative costs of information	54	54	54	54	54	54	54	54
B. Estimated total earnings benefits								
NPV change in total resources year 1 and beyond-until time horizon	1318				1318			
NPV change earnings for children as a result of changed stunting		503	675	589		503	675	589
C. Gain/cost ratio	**12.29**	**4.69**	**6.30**	**5.49**	**2.05**	**0.78**	**1.05**	**0.92**
D. Internal rate of return (IRR)	**218%**	**5.84%**	**3.78%**	**4.88%**	31%	-0.60%	-1.78%	-1.16%
E. Additional yearly benefits from age 2-16, in monthly food consumption terms (1 month = $11USD)								
1 month		9.15%	7.23%	8.28%		0.02%	-1.18%	-0.55%
2 months		13.7%	12.4%	13.1%		0.72%	-0.49%	0.15%
6 months		35.9%	35.9%	35.9%		4.50%	3.45%	4.01%
12 months		63.2%	63.2%	63.2%		12.6%	12.1%	12.4%

Notes: We analyze two potential scenarios; in scenario one, Columns 1-4, we assume the cash transfer as purely redistributive and therefore is not incorporated into the costs directly. In scenario two, Columns 5-8, we assume the transfers are a pure cost and are 100% incorporated into the costs. We assume in both scenarios that the administrative costs of cash transfers and the administrative costs of information are 10% of the cash transfer. All costs are presented in NPV terms with a 5% discount rate. To calculate the NPV change in total earnings we assume remaining expected productive life of new assets is 5 years after the transfers have stopped and take our ITT impact on net resources per month at Midline and Endline. We calculate a NPV with a social discount rate of 5%. To calculate the impact on child earnings we use the estimated coefficient from Hoddinot et al. [2013]. The authors estimate a 4% increase for males and 9% increase for females from a 1SD increase in HAZ at 24 months. We take our estimated ITT for males and females of .140 and .239 respectively to calculate the % impact on earnings of 2.60 and 3.52 for males and females respectively. To estimate life-cycle earnings we take the sample of parents and perform OLS regressions of earnings on 10-year age dummies (16-25, 26-35 etc.). This produces average earnings of males and females at different ages. We estimate the increase in earnings from these and then present them in NPV terms to calculate the IRR. For sensitivity analysis we calculate the IRR if we assume that there is some monetary gain for the children before the age of 16 from all the other benefits. We suppose increased yearly incomes in increments of the average monthly food consumption measured in our sample ($11USD) per year from the age of 2 to 16.

Figure 1: Timeline

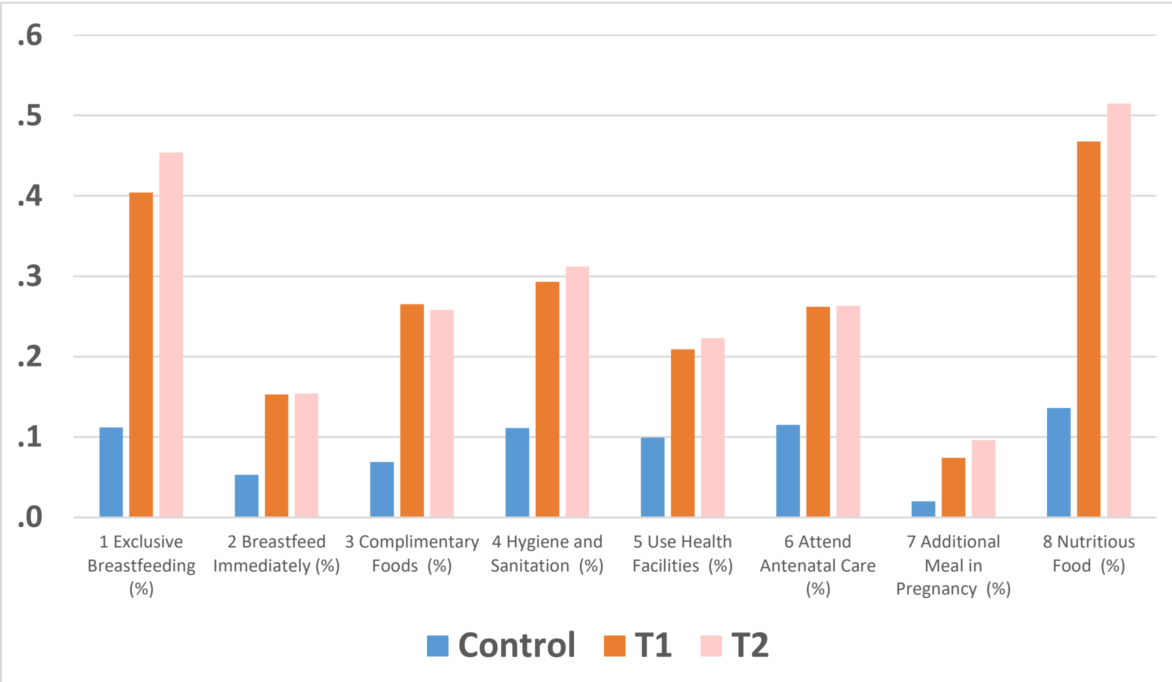


Notes: This depicts a timeline of the evaluation process for CDGP. The top part of the figure shows program implementation: when the registration began, when transfers began, when the program end was announced, and when it stopped enrolling new participants. The central part of the figure shows survey collection timings: when Baseline, Midline and Endline surveys were collected.

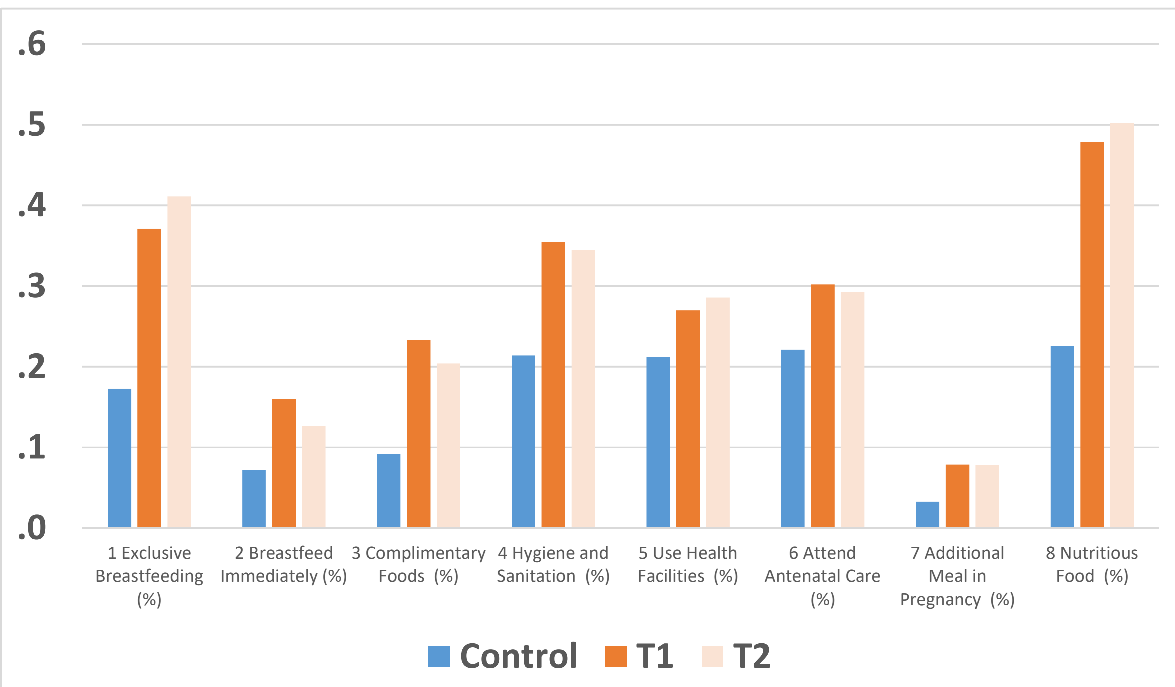
Figure 2: Recall of Key Messages, at Midline

Sample: Households with Pregnant Women at Baseline

Women



Husbands



Notes: This is based on women and their husbands in households with a pregnant woman at Baseline. It shows the proportion of treatment and control women and husband who recall the eight key messages at Midline. Recall is from any low intensity information channel (posters, radio, food demonstrations and health talks). Individuals are asked if they have been exposed to CDGP information from a particular information channel (and we repeat this for each channel). If the individual says yes to this, they are asked what messages do they recall from the information channel. If an individual was not exposed to any information channel, their recall of messages is set to zero.

Figure 3: HAZ Profile, Old Children at Baseline

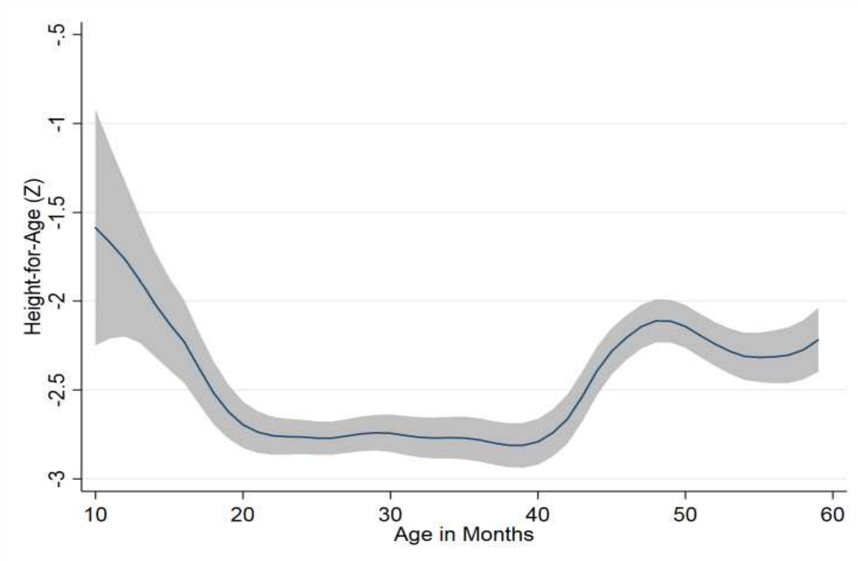
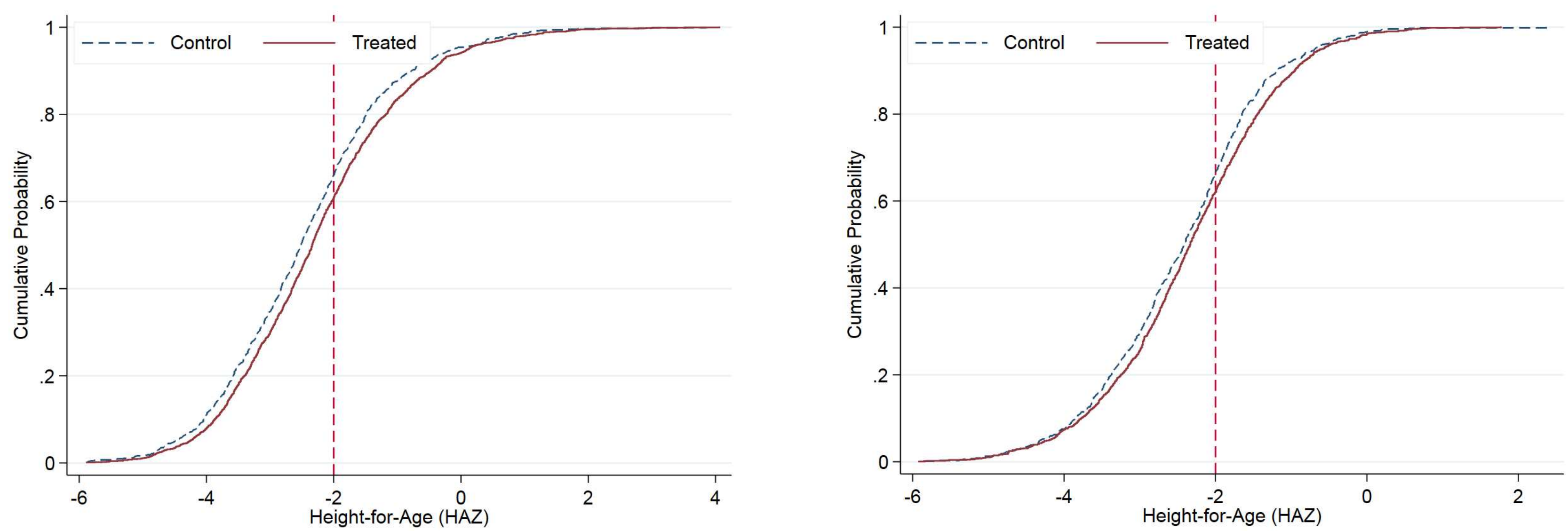


Figure 4: Distributional Impacts on Height-for-Age

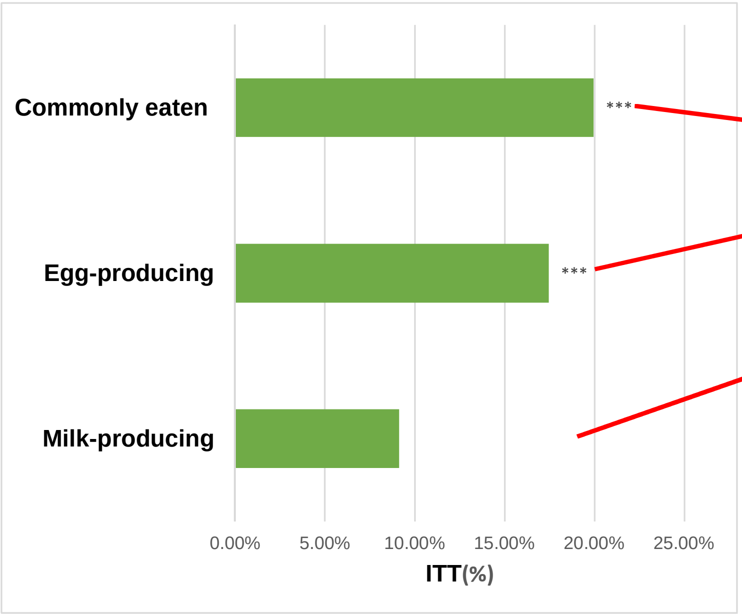
A. Midline **B. Endline**

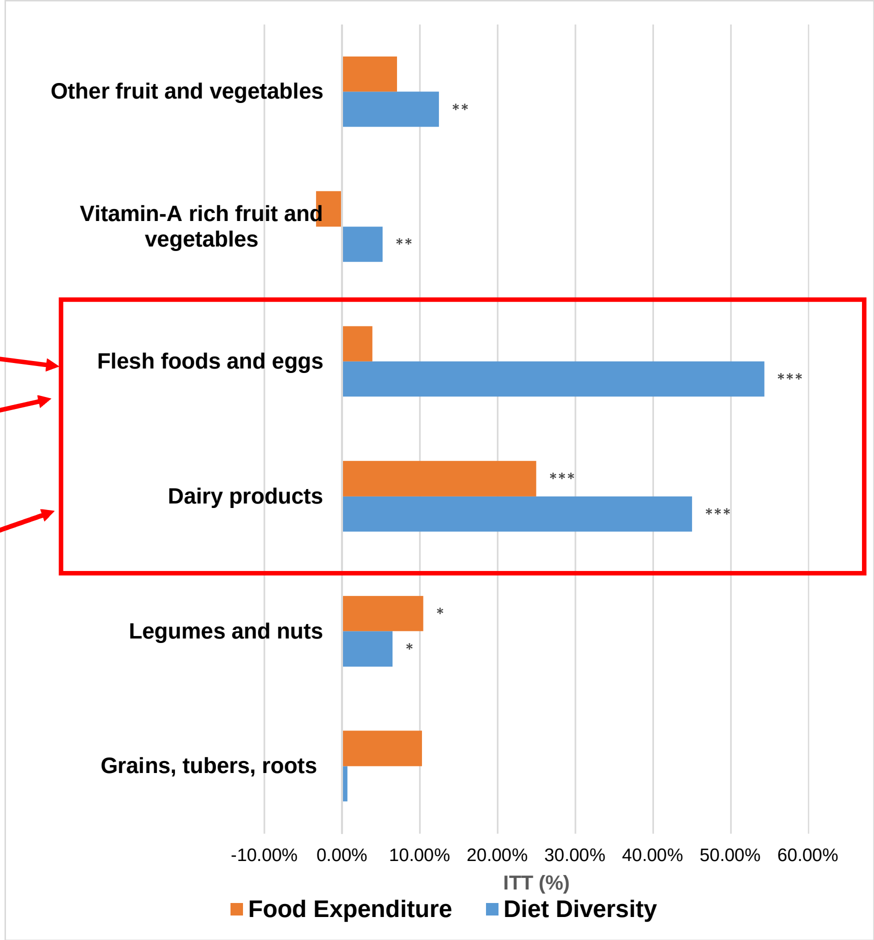


[illegible] at Baseli[illegible]ution of the HAZ score at Midline and Endline for the treatment and control group. A score to the left of the red dashed line indicates that the child is stunted (HAZ < -2).

Figure 5: Livestock Ownership, Dietary Diversity and Food Expenditures

Sample: Households with Pregnant Women at Baseline, In Utero Child (N=3688)

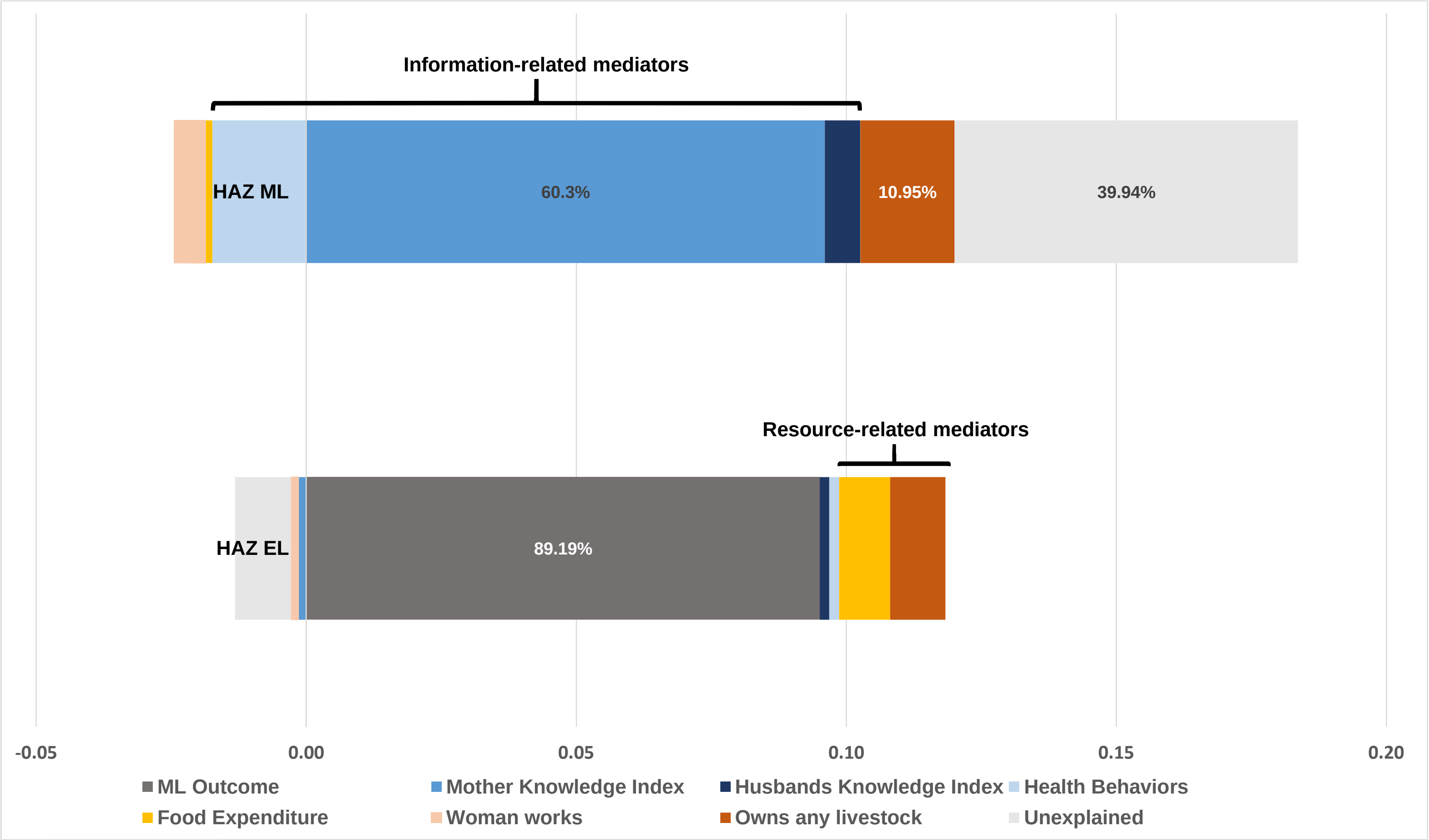




Notes: Significance levels: * (10%), ** (5%), ***(1%). There are two sets of bars in this figure showing ITT impacts on ownership of different types of livestock (left) and diet diversity / food expenditure (right) measured at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. For the ITTs on diet diversity in the right hand side Figure, in addition, we control for the gender of the child and the child's age in months. Standard errors are clustered at the village level throughout. On the left hand side (LHS) panel, the ITT effect is presented (in percentage points) and the lines represent 95% confidence intervals. On the LHS we group together animals owned by women into food producing groups. Milk producing animals include: female cow, goat, sheep. Commonly eaten animals include: cow, calf, sheep, goat. Egg producing animals include: chicken or guinea fowl. On the right hand side (RHS), the ITT estimate is then converted into a percentage impact over the Midline levels in Control villages. On the RHS, the diet diversity for the new child is obtained from a 24-hour food recall module administered to the child's mother or main carer. Each meal consumed in the day before the interview from waking up to bedtime is recorded, and each ingredient is coded into categories. On the right hand figure, all food expenditure categories are derived from 7-day recalls of expenditure. The top 1% of values are trimmed. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Figure 6: Mediation Analysis, HAZ

Sample: Households with Pregnant Women at Baseline, In Utero Child (N=3688)



Notes: This shows a decomposition of the ITT estimation on Height-for-age (Z) at Midline and Endline, following Gelbach [2016]. These ITTs are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. In addition, we control for the gender of the child and the child's age in months. The size of each horizontal bar corresponds to the total ITT effect, and each section of the bar is the part that can be explained by the individual mediator. Percentages of the largest sections indicate the proportion of the total size of the bar explained by that mediator. The unexplained proportion of the bar is that part of the ITT estimate that is not explained by the mediators. The mediator groups are defined as follows: Mother's knowledge index and Husband knowledge index (as previously defined), Health behavior index (as previously defined), Food expenditure, a dummy for if the woman works, a dummy for if the woman herself owns any livestock. The midline outcome is also included as mediator for the Endline estimates. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table A1: Information Components of the Intervention

A. Key Messages

Period	Message	Details
Prenatal	**Attend antenatal care**	Attend antenatal care at least four times during pregnancy.
	Eat one additional meal during pregnancy	Eat one extra small meal or 'snack' (extra food between meals) each day to provide energy and nutrients for you and your growing baby.
Perinatal	**Breastfeed immediately**	Start breast feeding your baby within the first 30 minutes of delivery. Colostrum is good for the baby.
	Breastfeed exclusively	Breastfeed your child exclusively until six months old. Do not give water, tinned milk, or any other food.
Postnatal	**Complementary feeding**	Introduce complimentary foods at six months of age while continuing to breastfeed. Breastfeed on demand and continue until two years of age. Gradually increase food variety as the child gets older.
	Hygiene and sanitation	Wash your hands after going to the toilet, cleaning baby who defecated, before and after feeding baby; wash baby's hands and face before feeding.
	Use health facilities	Take baby to health facility if you notice any of the following: fever, convulsion, refusing to eat, malnutrition, diarrhea.
	Nutritious food	Ensure you buy nutritious foods when you are buying food for your family.

B. Low- and High-intensity Channels of Message Delivery

T1: Low-Intensity	**Information and education posters**	Health and nutrition related posters are affixed in health facilities and village centers.
	Radio jingles / phone-in programs	Jingles are played regularly on local radio channels. Phone-in programs are one-hour shows in which CDGP staff and invited experts talk about one selected topic, and listeners can call in with questions.
	Friday preaching / Islamic school teachers	
	Health talks	Trained health workers come to the village and deliver a session on a selected topic, with the aid of information cards. Any village resident can attend these talks, irrespective of beneficiary status.
	Food demonstrations	CDGP trained staff delivers nutrition education about the benefits of different foods, and demonstrates how to prepare and cook nutritious meals for children and other household members.
	Voice messages	Pre-recorded messages are sent to beneficiaries' program phones to reinforce key messages.
T2: High-Intensity	**Infact and Young Child Feeding (ICYF) support groups**	Groups are formed within communities to support beneficiaries, under the supervision and facilitation of community volunteers and health extension workers. The recommended size is 12-15 people, meeting once a month. They are also offered to men.
	One-on-one counselling	Beneficiaries and their husbands can consult community volunteers on an `as needed' basis to receive specific information and training.

Notes: Panel A lists the eight key messages around which the behavior change communication component of CDGP was built. Panel B details the channels by which these key messages were delivered to beneficiaries in treated villages.

Table A2: Attrition

Dependent variable: attrit from sample (0/1)

Standard errors in parentheses clustered by village

	Pregnant Woman at Baseline			Husband	Old Child	New Child
Period: Baseline to	**(1) EL**	**(2) EL**	**(3) EL**	**(4) EL**	**(5) ML**	**(6) EL**
T1	.012	.009	.032	.058	.049	.085
	(.010)	(.010)	(.074)	(.078)	(.091)	(.089)
T2	.015	.012	.081	.098	.043	.041
	(.011)	(.011)	(.073)	(.077)	(.097)	(.093)
Village insecure at midline	.028*	.019	.020	.017	.887***	
	(.013)	(.017)	(.019)	(.017)	(.017)	
Village insecure at endline	.892***	.876***	.897***	.883***		.832***
	(.010)	(.013)	(.017)	(.017)		(.032)
T1 * Village insecure at midline					-.024	
					(.025)	
T2 * Village insecure at midline					-.032	
					(.026)	
T1 * Village insecure at endline			-.032	-.029		-.023
			(.018)	(.019)		(.035)
T2 * Village insecure at endline			-.027	-.026		-.007
			(.019)	(.020)		(.038)
Randomization Strata	Yes	Yes	Yes	Yes	Yes	Yes
Attrition rate	.227	.227	.227	.241	.198	.203
Joint p-value on interactions	-	-	.503	.740	.781	.279
Observations	3688	3688	3688	3688	2597	2719

Notes: Significance levels: * (10%), ** (5%), ***(1%). Each Column presents estimates using a linear probability model where the dependent variable is if the individual subject attrits and the independent variables are a varying set of treatment indicators, baseline covariates and interactions. Attrition takes the value of one if the subject surveyed at Baseline (or Midline if the New Child) was not surveyed at Endline (except for attrition of the Old Child, which is measured at Midline). The sample in Columns 1 to 3 are women pregnant at Baseline. In Column 4, the sample is husbands of women who were pregnant at Baseline. In Column 5, the sample is the Old Child in households where the woman was pregnant at Baseline. In Column 6, the sample is the New Child in households where the woman was pregnant at Baseline. All Columns include treatment status and village insecurity status, at Midline and Endline. Column 2 adds controls for Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. All other Columns further add interactions between the program indicators and the covariates as well as interaction between security and treatment status. Column 5 does not include insecure at Endline as the Old Child is not surveyed then. Column 6 does not include insecure at Midline as the New Child is only followed from Midline onwards. At the foot of Columns 3 onwards, we report the p-value on the null on the joint hypothesis test that all interaction terms are zero.

Table A3: Exposure to Low- and High-Intensity Channels of Information Dissemination

Sample: Households with pregnant women at baseline (N=3688)
Means, p-values in brackets

	Wife				Husband				Wife = Husband		
	(1) Control	(2) T1	(3) T2	(2) = (3)	(4) Control	(5) T1	(6) T2	(5) = (6)	(1) = (4)	(2) = (5)	(3) = (6)
Panel A: Low-intensity channels											
At least one (%)	62.7	87.6	91.4	[.301]	67.8	91.8	92.4	[.979]	[.139]	[.184]	[.752]
All (%)	2.28	21.7	25.2	[.291]	.956	4.29	6.67	[.026]	[.069]	[.000]	[.000]
Panel B: High-intensity channels											
None (%)	80.4	58.0	45.5	[.005]	94.9	79.5	77.6	[.364]	[.000]	[.000]	[.000]
All (%)	1.60	11.8	18.7	[.003]	1.43	7.47	5.79	[.517]	[.083]	[.009]	[.000]
Support group	18.9	33.5	43.6	[.024]	4.00	14.7	16.8	[.345]	[.000]	[.000]	[.000]
Says 1:1 counselling available (%)	24.2	62.9	63.2	[.949]	13.2	49.0	48.5	[.929]	[.009]	[.000]	[.000]
If yes: tried to obtain 1:1 counselling	30.4	32.2	47.5	[.000]	23.7	27.0	34.4	[.101]	[.321]	[.261]	[.014]
If yes: obtained 1:1 counselling (%)	95.3	89.7	91.4	[.657]	97.9	94.8	95.5	[.946]	[.352]	[.145]	[.234]

Notes: Column 1-3 show the means of sampled women's exposure to information channels in the Control and each Treatment groups. Columns 4-6 show the corresponding means for husbands. Column 2=3 the p-values that test the hypothesis that the estimated effects are equal between treatment arms for women. Column 5=6 report the p-values for husbands. Columns 1=4, 2=5 and 3=6 show the p-values that test the hypothesis that the estimated effects are equal between wife and husband, within each treatment arm (Control, T1 and T2). P-values are derived from an OLS regression that controls for randomization strata, and clusters standard errors by village. In Panel A, low-intensity channels include posters, radio, attending food demonstrations and attending health talks. In Panel B, high-intensity channels include 1:1 counselling and support groups. The answers to 1:1 counselling are answered sequentially, so that the next answer is given that the respondent answered yes in the previous question

Table A4: Recall of Messages at Midline from Low Intensity Channels

Sample: Households with pregnant women at baseline (N=3688)
Means, p-values in brackets

	Wife				Husband				Wife = Husband		
	(1) Control	(2) T1	(3) T2	(2) = (3)	(4) Control	(5) T1	(6) T2	(5) = (6)	(1) = (4)	(2) = (5)	(3) = (6)
All (%)	.001	.003	.004	[.762]	.003	.005	.004	[.696]	[.151]	[.415]	[1.00]
At least one (%)	.290	.702	.723	[.271]	.467	.767	.770	[.727]	[.000]	[.001]	[.001]
None (%)	.710	.298	.277	[.271]	.533	.233	.230	[.727]	[.000]	[.001]	[.001]
Number (%)	.716	2.13	2.28	[.091]	1.24	2.25	2.25	[.643]	[.000]	[.066]	[.550]
1 Exclusive Breastfeeding (%)	.112	.404	.454	[.028]	.173	.371	.411	[.049]	[.000]	[.073]	[.007]
2 Breastfeed Immediately (%)	.053	.153	.154	[.842]	.072	.160	.127	[.113]	[.016]	[.531]	[.001]
3 Complimentary Foods (%)	.069	.265	.258	[.941]	.092	.233	.204	[.146]	[.017]	[.036]	[.000]
4 Hygiene and Sanitation (%)	.111	.293	.312	[.396]	.214	.355	.345	[.861]	[.000]	[.000]	[.005]
5 Use Health Facilities (%)	.099	.209	.223	[.554]	.212	.270	.286	[.43]	[.000]	[.000]	[.000]
6 Attend Antenatal Care (%)	.115	.262	.263	[.525]	.221	.302	.293	[.934]	[.000]	[.000]	[.030]
7 Additional Meal in Pregnancy (%)	.020	.074	.096	[.051]	.033	.079	.078	[.786]	[.025]	[.414]	[.020]
8 Nutritious Food (%)	.136	.468	.515	[.024]	.226	.479	.502	[.165]	[.000]	[.459]	[.306]

Notes: Column 1-3 show the means of sampled women's recall of messages from low-intensity channels in the Control and each Treatment groups. Columns 4-6 show the corresponding means for husbands. Column 2=3 the p-values that test the hypothesis that the estimated effects are equal between treatment arms for women. Column 5=6 report the p-values for husbands. Columns 1=4, 2=5 and 3=6 show the p-values that test the hypothesis that the estimated effects are equal between wife and husband, within each treatment arm (Control, T1 and T2). P-values are derived from an OLS regression that controls for randomization strata, and clusters standard errors by village. Low-intensity channels include posters, radio, attending food demonstrations and attending health talks.

Table A5: Anthropometrics, Age Adjustments

Sample: Households with pregnant women at baseline (N=3688)

Standard errors in parentheses clustered by village, p-values in brackets

	Age-Adjusted ITT, Midline			Age-Adjusted ITT, Endline			(1) = (4)	(2) = (5)	(3) = (6)
Age control:	(1) NP	(2) Cubic	(3) CF	(4) NP	(5) Cubic	(6) CF			
Height-for-Age (HAZ)	.133**	.141**	.166***	.110*	.124**	.157	[.702]	[.755]	[.938]
	(.064)	(.058)	.058	(.061)	(.059)	(.116)			
Stunted (HAZ < -2)	-2.57	-3.26	-3.55	-4.86*	-5.21**	-6.35	[.389]	[.434]	[.598]
	(2.47)	(2.17)	(2.328)	(2.62)	(2.60)	(4.773)			
Severely stunted (HAZ < -3)	-3.74	-3.47*	-4.62**	-3.91*	-4.71**	-5.12	[.943]	[.570]	[.921]
	(2.43)	(2.05)	(2.120)	(2.15)	(2.17)	(4.647)			
Weight-for-Age (WAZ)	.021	.016	.031	.050	.041	.067	[.610]	[.650]	[.769]
	(.063)	(.055)	(.055)	(.055)	(.055)	(.112)			
Weight-for-height (WHZ)	-.116**	-.097*	-.091**	-.046	-.057	-.048	[.286]	[.538]	[.730]
	(.052)	(.050)	(.047)	(.056)	(.057)	(.115)			
Middle upper arm circumference (MUAC)	4.39	3.72	3.65	1.29*	1.35*	-1.52	[.132]	[.230]	[.426]
	(3.23)	(2.97)	(2.314)	(.702)	(.705)	(6.07)			
Malnourished (MUAC < 125mm)	.103	.756	.341	-.838	-1.17*	-.597	[.602]	[.269]	[.801]
	(1.76)	(1.65)	(1.04)	(.587)	(.624)	(3.56)			

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) value in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships, and the gender of the new child. Columns 1 and 4 control for the age non-parametrically using dummies for different age ranges of the New Child. The age dummies (in months) are: 14-20, 21-27 at Midline and 21-27, 28-33, 34-39, 40-45, 46-51 at Endline. Columns 2 and 5 control for age using a cubic in age in months. Columns 3 and 6 present control function estimations. The estimations control for age with the same age dummies as in Columns 1 and 4 and in addition control functions are estimated as follows: in the first stage, the age of the child is regressed on all covariates and the date of interview (the exogenous instrument); residuals from the first stage are then squared and cubed and included in the regression for the outcome. In the control function specifications in Columns 3 and 6, standard errors are computed by bootstrap with 1,000 repetitions. In all other Columns, standard errors are clustered at the village level. Stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -2 standard deviations of the WHO defined guidelines [WHO 2009]. Severely stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -3 standard deviations of the WHO defined guidelines. Wasted is defined as being below -2 standard deviation below weight-for-height (WHZ) WHO defined guidelines. Malnourished is a dummy indicating children with a MUAC of less than 125mm.

Table A6: Anthropometric Impacts by Gender

Sample: Households with Pregnant Women at Baseline (N=3688)

Standard deviation in braces

Standard errors in parentheses clustered by village, p-values in brackets

	Boys				Girls				Girls = Boys	
	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)	(4) Control Mean	(5) ITT, Midline	(6) ITT, Endline	(5) = (6)	ML	EL
Height-for-Age (HAZ)	-2.46	.239*	.086	[.055]	-2.13	.140*	.158**	[.836]	[.337]	[.459]
	{1.41}	(.088)	(.073)		{1.33}	(.084)	(.075)			
Stunted (HAZ < -2)	68.3	-4.50	-5.27*	[.832]	63.7	-5.79*	-4.33	[.703]	[.792	[.765]
		(3.08)	(3.15)			(3.28)	(3.37)			
Severely stunted (HAZ < -3)	38.8	-4.45	-3.05	[.663]	30.2	-4.53*	-5.41*	[.767]	[.998]	[.529]
		(3.06)	(2.59)			(2.69)	(2.85)			
Weight-for-Age (WAZ)	-1.79	.069	-.018	[.181]	-1.66	-.026	.130*	[.053]	[.335]	[.083]
	{1.18}	(.078)	(.067)		{1.22}	(.079)	(.069)			
Weight-for-height (WHZ)	-.652	-.090	-.135	[.727]	-.594	-.158**	.044	[.027]	[.491	[.068]
	{1.14}	(.064)	(.968)		{1.13}	(.072)	(.072)			
Wasted (WHZ < -2)	12.8	2.25	-.413	[.178]	9.32	3.95**	-1.35	[.011]	[.489]	[.456]
		(1.69)	(.965)			(1.80)	(1.18)			
Middle upper arm circumference (MUAC)	137	-.217	.513	[.451]	133	-.606	1.41	[.052]	[.321]	[.516]
	{13.0}	(.866)	(.854)		{12.9}	(.812)	(.897)			
Malnourished (MUAC < 125mm)	14.5	1.33	-.060	[.499]	21.1	.608	-1.64*	[.375]	[.849]	[.256]
		(2.07)	(.771)			(2.46)	(.985)			

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) value in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, and a dummy for polygamous relationships. The estimates also control for the age non-parametrically using dummies for different age ranges of the New Child. The age dummies (in months) are: 14-20, 21-27 at Midline and 21-27, 28-33, 34-39, 40-45, 46-51 at Endline. Standard errors are clustered at the village level throughout. Stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -2 standard deviations of the WHO defined guidelines [WHO 2009]. Severely stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -3 standard deviations of the WHO defined guidelines. Severely stunted is a dummy indicating children with height-for-age-z-score (HAZ) under -3 standard deviations of the WHO defined guidelines. Wasted is defined as being below -2 standard deviation below weight-for-height (WHZ) WHO defined guidelines. Malnourished is a dummy indicating children with a MUAC of less than 125mm.

Table A7: Child Development and Maternal Time with Child

Sample: Households with pregnant women at baseline (N=3688)

Column 1: Standard deviation in braces

Columns 2 to 5: Standard errors in parentheses clustered by village

	Midline		Endline	
	(1) Control Mean	**(2) ITT**	**(3) ITT**	**(2) = (3)**
Panel A: Child Development				
Communication Skills (Z)	.000	.133**	.047	[.226]
	{1.00}	(.055)	(.058)	
		[.291]	[.803]	
Low Communication Skills (%)	68.0	-4.43*	1.93	[.025]
		(2.34)	(1.40)	
		[.549]	[.803]	
Gross Motor Skills (Z)	.000	.078	.074	[.952]
	{1.00}	(.054)	(.059)	
		[.780]	[.545]	
Low Gross Motor Skills (%)	60.0	-3.82	2.25	[.065]
		(2.75)	(1.77)	
		[.803]	[.803]	
Personal-Social Skills (Z)			-.097	
			(.064)	
			[.778]	
Low Personal-Social Skills (%)			2.55	
			(2.76)	
			[.803]	
Panel B: Daily Time Mother Spent Playing with New Child				
< 2 hours (%)	72.6	-4.96*		
		(2.68)		
		[.042]		
2-5 hours (%)	21.2	.836		
		(2.27)		
		[.746]		
> 5 hours (%)	4.86	4.73***		
		(1.46)		
		[.001]		

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) value in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships, and the gender of the new child. Standard errors are clustered at the village level throughout. Panel A reports child development scores that are obtained from the Age and Stages Questionnaire (ASQ). The survey included questions on communication and gross motor skills at Midline and in addition included personal-social skills at Endline. The standardized test scores are standardized for each age so that the control mean at any given age (in months) is 0. Low scores is a dummy indicating the child's score falls below the 'normal range'.

Table A8: Household Dietary Diversity and Food Expenditures, by Food Group

Sample: Households with pregnant women at baseline (N=3688)

Standard deviation in braces

Standard errors in parentheses clustered by village

	Dietary Diversity (%)				Food Expenditures (USD)			
	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)	(4) Control Mean	(5) ITT, Midline	(6) ITT, Endline	(5) = (6)
MDD 1: Grains, tubers, roots	93.3	2.05**	0.67	[.209]	51.2	6.38	5.27	[.833]
		(0.94)	(0.52)		{73.0}	(4.11)	(3.65)	
MDD 2: Legumes and nuts	60.9	3.08	3.97*	[.790]	15.2	1.16	1.59*	[.769]
		(2.30)	(2.21)		{24.9}	(1.27)	(0.89)	
MDD 3: Dairy products	26.6	15.43***	11.98***	[.286]	5.2	1.42***	1.30***	[.823]
		(2.38)	(2.88)		{9.1}	(0.46)	(0.36)	
MDD 4 and 5: Flesh foods and eggs	19.2	8.12***	10.43***	[.401]	30.9	5.98***	1.21	[.050]
	{39.4}	(2.23)	(2.58)		{37.8}	(2.02)	(1.78)	
MDD 6: Vitamin-A rich fruit and vegetables	74.4	2.03	3.91**	[.439]	2.7	0.75***	-0.09	[.010]
		(1.94)	(1.55)		{4.8}	(0.27)	(0.24)	
MDD 7: Other fruit and vegetables	47.0	5.50**	5.87**	[.907]	13.1	0.98	0.93	[.958]
		(2.53)	(2.45)		{16.9}	(0.92)	(0.69)	
Other: Oil, butter and condiments					26.4	0.04	2.41**	[.106]
					{25.5}	(1.30)	(0.94)	
Other: Sugary items, drinks					5.2	0.84*	0.83***	[.984]
					{8.7}	(0.49)	(0.31)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. In addition, we control for the gender of the child and the child's age in months. Standard errors are clustered at the village level throughout. The diet diversity index and the food groups are obtained from a 24-hour food recall module administered to the child's mother or main carer. Each meal consumed in the day before the interview from waking up to bedtime is recorded, and each ingredient is coded into categories. Columns 4-6 present food expenditures of matching MDD food groups plus two additional categories (oil/butter/condiments and sugary items/drinks). All expenditure categories are derived from 7-day recalls of expenditure, with the top 1% of values being trimmed. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table A9: Livestock Ownership

Sample: Households with pregnant women at baseline (N=3688)

Standard deviation in braces

Standard errors in parentheses clustered by village

	Household				Wife			
	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)	(4) Control Mean	(5) ITT, Midline	(6) ITT, Endline	(5) = (6)
Owns any animals (%)	89.8	0.255	4.97***	[.011]	78.3	6.56***	12.1***	[.016]
		(1.37)	(1.53)			(1.94)	(2.12)	
Any goat (%)	71.2	0.62	6.61***	[.024]	56.9	7.38***	15.0***	[.005]
		(2.24)	(2.30)			(2.50)	(2.54)	
Any chicken (%)	60.9	0.66	3.03	[.468]	38.8	6.52***	8.78***	[.421]
		(2.66)	(2.49)			(2.43)	(2.74)	
Any sheep (%)	55.9	-1.71	5.64**	[.003]	33.1	2.50	8.09***	[.014]
		(2.44)	(2.29)			(2.05)	(2.23)	
Any camel (%)	4.73	0.082	-0.778	[.396]	.000	.089	.445*	[.152]
		(.869)	(.968)			(.067)	(.245)	
Any cow/bull (%)	36.5	-.550	4.07	[.053]	4.37	.027	1.13	[.425]
		(2.42)	(2.64)			(.859)	(1.29)	
Any donkey (%)	3.09	-0.142	-0.478	[.684]	.299	-.260	.429**	[.005]
		(.929)	(.740)			(.173)	(0.211)	
Any guinea fowl (%)	16.5	-2.23	-0.45	[.457]	4.86	-.122	1.08	[.355]
		(1.96)	(2.44)			(.823)	(1.16)	
Any calf (%)	13.6	1.63	2.87	[.568]	3.08	1.02	1.73*	[.500]
		(1.66)	(2.03)			(.801)	(1.04)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Columns 1-3 report results for household ownership of livestock, and Columns 4-6 report results for wife's ownership of livestock. Columns 1 and 4 shows the mean (and standard deviation for continuous outcomes) values in Control households at Baseline. Columns 2 and 5 reports ITT estimates at Midline, and Columns 3 and 6 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout.

Table A10: Prices of Livestock, $US [PPP]

Sample: All households

Means, standard deviation in braces

	Price Paid to Purchase			Revenue from Selling		
	(1) Obs	**(2) Mean (SD)**	**(3) Median**	**(4) Obs**	**(5) Mean (SD)**	**(6) Median**
Female Sheep	167	83.9	80.5	272	245	132
		{31.6}			{596}	
Male Sheep	324	125	121	404	330	201
		{56.4}			{477}	
Female Goat	238	54.2	50.3	456	120	80.5
		{20.6}			{142}	
Male Goat	147	66.1	60.4	231	117	80.5
		{34.7}			{115}	
Chicken				143	49.3	22.1
					{93.7}	

Notes: The sample for this table is all households interviewed, irrespective of whether the women was pregnant or not at baseline. Columns 1 and 4 report the number of observations used to construct each price estimate. Columns 2 and 5 report the mean price (and standard deviation) and Columns 3 and 6 report the median price. Columns 1-3 report details on the price paid to purchase different animals. Columns 4-6 report the revenue from selling the animals reported. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table A11: Saving and Borrowing, Detailed

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Panel A: Saving				
Saving money, including In kind (%)	42.0	1.71	6.73**	[.240]
		(2.61)	(2.82)	
Panel B: Borrowing				
Any household member borrowing (%)	23.4	-3.01	-7.46***	[.126]
		(2.33)	(2.41)	
Any household member failed to borrow (%)	8.1	-1.43	-1.27	[.961]
		(2.27)	(2.23)	
Panel C: Lending				
Any member of household providing loans (%)	14.4	-3.90*	1.94	[.044]
		(2.09)	(2.01)	
Total value of loans	1.3	-0.57	-2.01	[.468]
	{17.7}	(0.99)	(1.65)	

Notes: Significance levels: * (10%), ** (5%), ***(1%). In Panels A, B and C, Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Baseline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout. For continuous monetary outcomes, values above the 99th percentile are set to missing. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table A12: Mediation Analysis

Sample: Households with pregnant women at baseline (N=3688)
Standard errors in parentheses clustered by village

	Outcome: Height-for-age (HAZ)				Outcome: NC Child Health Index			
	Midline		Endline		Midline		Endline	
	(1) Coefficient	(2) Fraction Mediated	(3) Coefficient	(4) Fraction Mediated	(5) Coefficient	(6) Fraction Mediated	(7) Coefficient	(8) Fraction Mediated
Unconditional Effect (ITT)	.159		.107		.192		.303	
	(.074)		(.058)		(.057)		(.053)	
Conditional Effect (ITT)	.064		-.009		.166		.274	
	(.077)		(.053)		(.061)		(.057)	
Mediated by:								
Mother knowledge index	.096	60.31	-.001	-1.35	.046	24.10	.029	9.51
	(.035)		(.022)		(.023)		(.022)	
Husband knowledge index	.007	4.13	.002	1.64	-.007	-3.47	-.003	-0.91
	(.005)		(.003)		(.005)		(.003)	
Health behaviors index	-.017	-10.94	.002	1.74	-.014	-7.24	-.010	-3.44
	(.007)		(.005)		(.005)		(.006)	
Food expenditure	-.001	-0.78	.009	8.86	-.002	-1.20	-.002	-0.56
	(.006)		(.008)		(.005)		(.007)	
Woman works	-.006	-3.61	-.001	-1.32	.001	0.41	-.001	-0.43
	(.006)		(.007)		(.004)		(.007)	
Owns any livestock	.017	10.95	.010	9.60	.002	1.29	-.014	-4.56
	(.015)		(.008)		(.012)		(.008)	
ML outcome			.095	89.19			.030	9.78
			(.034)				(.009)	
Total mediation	.096	60.06	.116	108.35	.027	13.89	.028	9.40
	(.033)		(.041)		(.023)		(.025)	

Notes: This shows a decomposition of the ITT estimation on Height-for-age (Z)) at Midline and Endline (Columns 1-3, and for the New Child Health Index at Midline and Endline (Columns 5-7), following Gelbach [2016]. These ITTs are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. In addition, we control for the gender of the child and the child's age in months. The mediator groups are defined as follows: Mother's knowledge index and Husband knowledge index (as previously defined), Health behaviour index (as previously defined), Food expenditure, a dummy for if the woman works, a dummy for if the woman herself owns any livestock. The midline outcome is also included as mediator for the Endline estimates. All monetary amounts are converted from Nigerian Naira to PPP US dollars at the 2014 rate.

Table A13: Maternal Health

Sample: Households with pregnant women at baseline (N=3688)
Standard deviation in braces
Standard errors in parentheses clustered by village
Romano and Wolf [2016] Bootstrap p-values in square brackets below standard errors

	(1) Control Mean	(2) ITT, Midline	(3) ITT, Endline	(2) = (3)
Weight	49.8	-.382	-.412	[.897]
	{7.33}	(.258)	(.289)	
		[.999]	[.999]	
Height	157	-.044	.120	[.077]
	{5.56}	(.095)	(.103)	
		[.995]	[.935]	
BMI	20.1	-.137	-.181	[.631]
	{2.63}	(.103)	(.114)	
		[.999]	[.999]	
BMI: Thin	.279	.021	.043*	[.317]
		(.024)	(.022)	
		[.999]	[.922]	
BMI: Normal	.665	-.023	-.044*	[.386]
		(.028)	(.025)	
		[.996]	[.457]	
BMI: Overweight	.056	.004	.003	[.978]
		(.013)	(.013)	
		[.999]	[.868]	
Mid-upper Arm Circumference	253	-.996	.487	[.167]
	{25.0}	(.990)	(1.28)	
		[.986]	[.997]	
Moderately Malnourished	.071	.019	.008	[.416]
		(.013)	(.013)	
		[.974]	[.999]	
Severely Malnourished	.000	.001	.005**	[.172]
		(.001)	(.002)	
		[.234]	[.997]	

Notes: Significance levels: * (10%), ** (5%), ***(1%). Column 1 shows the mean (and standard deviation for continuous outcomes) values in Control households at Midline. Column 2 reports ITT estimates at Midline, and Column 3 reports ITT estimates at Endline. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. SEs are clustered at the village level. We also report p-values adjusted for multiple testing in square brackets below the standard errors for certain outcomes. These are computed using the step-down procedure discussed in Romano and Wolf [2016], with 1,000 bootstrap replications. There are nine outcomes being simultaneously tested at midline and endline. Therefore, the p-values are adjusted for testing on 18 hypotheses.

Table A14: Impacts on Older and Younger Siblings

Sample: Households with pregnant women at baseline (N=3688)

Standard errors in parentheses clustered by village

	New Child	Old Child		Endline Child	
		Older Sibling of New Child		Younger Sibling of New Child	
	ITT, Midline	(1) Control Mean	(2) ITT, Midline	(3) Control Mean, Endline	(4) ITT, Endline
Panel A: Height and Stunting					
Height-for-Age (HAZ)	.198***	-2.48	-.091**	-1.72	.016
	(.068)	{1.47}	(.050)	{1.63}	(.079)
Stunted (HAZ < -2) (%)	-5.22**	57.9	2.78	45.8	2.38
	(2.43)		(3.17)		(2.22)
Severely stunted (HAZ < -3) (%)	-4.77**	22.2	1.76	21.3	-.359
	(2.21)		(2.56)		(1.90)
Panel B: Health					
Health Outcome Index	.209***	.000	.123**	.000	.214***
	(.052)	{1.00}	(.054)	{1.00}	(.056)
Panel C: Health Practices and Behaviors					
Antenatal Practices Index	.841***			.000	.223***
	(.085)			{1.00}	(.064)
Health Behaviors Index	.173***	.000	.238***	-.018	.214***
	(.057)	{1.00}	(.054)	(1.00)	(.056)
Panel D: Diet					
Minimum Dietary Diversity Indicator	.364***	2.62	.235***	2.59	.623***
	(.071)	{1.09}	(.065)	{1.68}	(.099)

Notes: Significance levels: * (10%), ** (5%), ***(1%). The non-numbered first Column reports the ITT estimates at Midline for the new child, that have been reported in previous tables. Column 1 reports the mean (and standard deviation for continuous outcomes) for control households at Baseline for the Old Child. Column 2 reports ITT estimates at Midline for the old child. Column 3 reports the mean values in control households at Endline for the End Child. Column 4 reports ITT estimates at Endline for the end child. Each ITT is estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout. Panel A presents non-parametric age-adjusted estimates for the End Child. The specification controls for the age non-parametrically using dummies for different age ranges of the End Child. The age dummies (in months) are: 0-5, 6-11, 12-17, 18-23, 24-29. The index variables are computed using the methodology in Anderson (2008), and are standardized to have mean zero and variance one in the control group. The Health Outcome Index includes the following components: a dummy variable that takes the value of 1 if the child has not been ill in the last month and a dummy variable that takes the value of 1 if the child has not had diarrhea in the past two weeks. The Health behavior index is similarly constructed from the following question components: Has the child been given deworming medicine in the past 6 months? Has the child received all basic vaccinations? The Antenatal Practices Index includes the following components: Had any antenatal care for the child? A doctor/nurse/midwife/CHEW assisted the birth of the child. The child was put the breast immediately. The child was appropriately breastfed. The Minimum Dietary Diversity Indicator sums the number of food groups the child has received from the following 7 food groups: 1. Grains, roots and tubers, 2. Legumes and nuts, 3. Dairy products, 4. Flesh foods, 5. Eggs, 6. Vitamin-A rich fruits and vegetables, 7. Other fruits and vegetables.

Table A15: Impacts on Older and Younger Siblings, Detailed

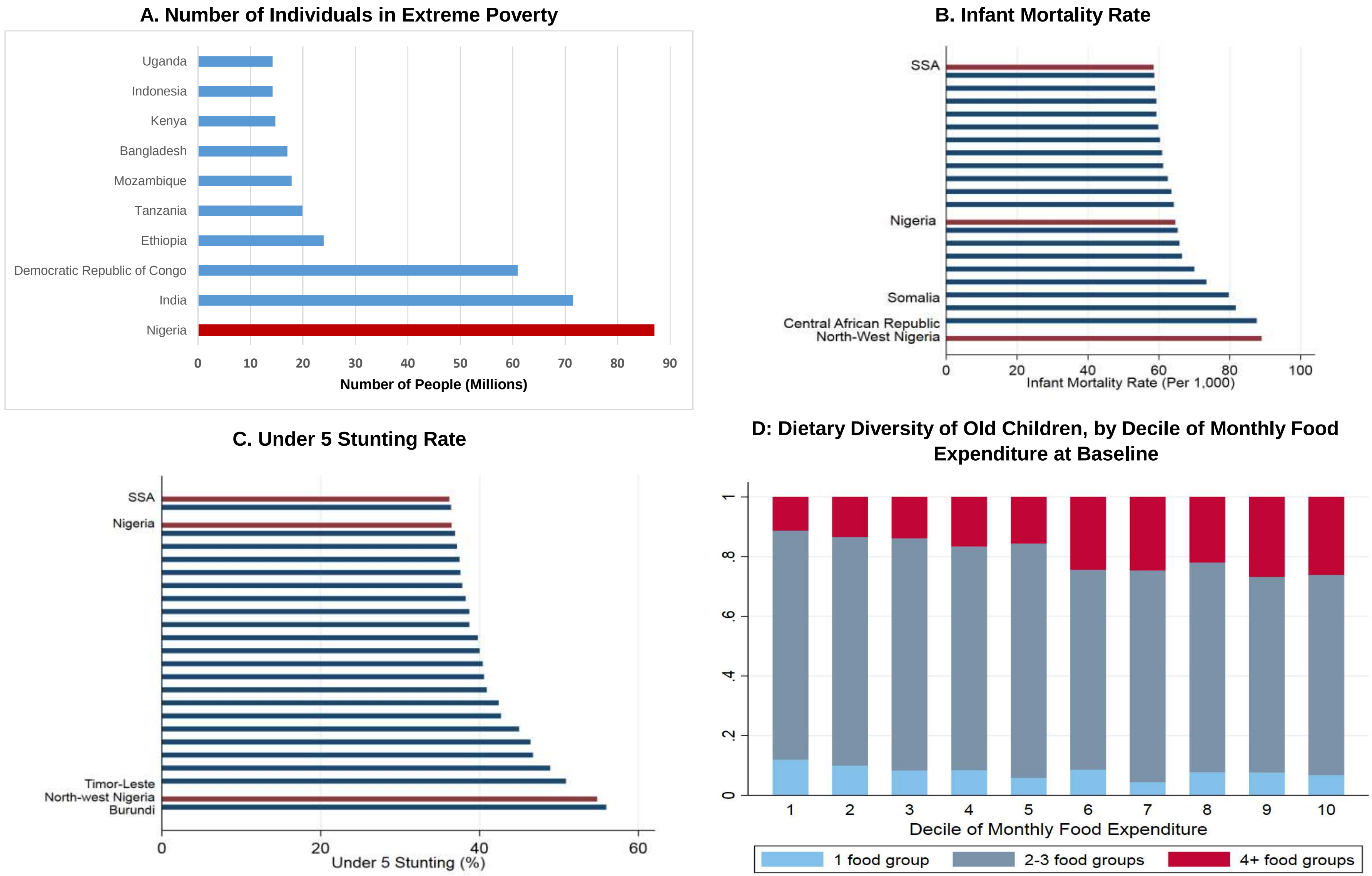
Sample: Households with pregnant women at baseline (N=3688)

Standard errors in parentheses clustered by village

	New Child	Old Child		Endline Child	
		Older Sibling of New Child		Younger Sibling of New Child	
	ITT, Midline	(1) Control Mean	(2) ITT, Midline	(3) Control Mean	(4) ITT, Endline
Panel A: Antenatal Practices					
Had antenatal care (%)	16.8***			69.6	17.0***
	(3.26)			(46.0)	(3.30)
Put to breast immediately (%)	4.75**			67.2	20.1***
	(2.00)			(47.0)	(2.73)
Born at health facility (%)	26.3***			15.1	11.0***
	(3.14)			(35.9)	(2.38)
Exclusively breastfed for 6 months (%)	29.0***			68.5	6.18**
	(2.87)			(46.5)	(2.52)
Panel B: Health Behaviors					
If had diarrhea in past two weeks:					
Anyone sought advice/treatment (%)	6.88**	80.4	6.71*	81.1	1.15
	(3.02)		(4.06)		(3.62)
Given ORS for diarrhea (%)	10.3**	45.7	10.6**	43.9	14.0***
	(4.03)		(5.30)		(4.93)
Given deworming medication in past 6 months (%)	8.02***	21.2	11.6***	21.0	10.5***
	(2.43)		(2.51)		(2.13)

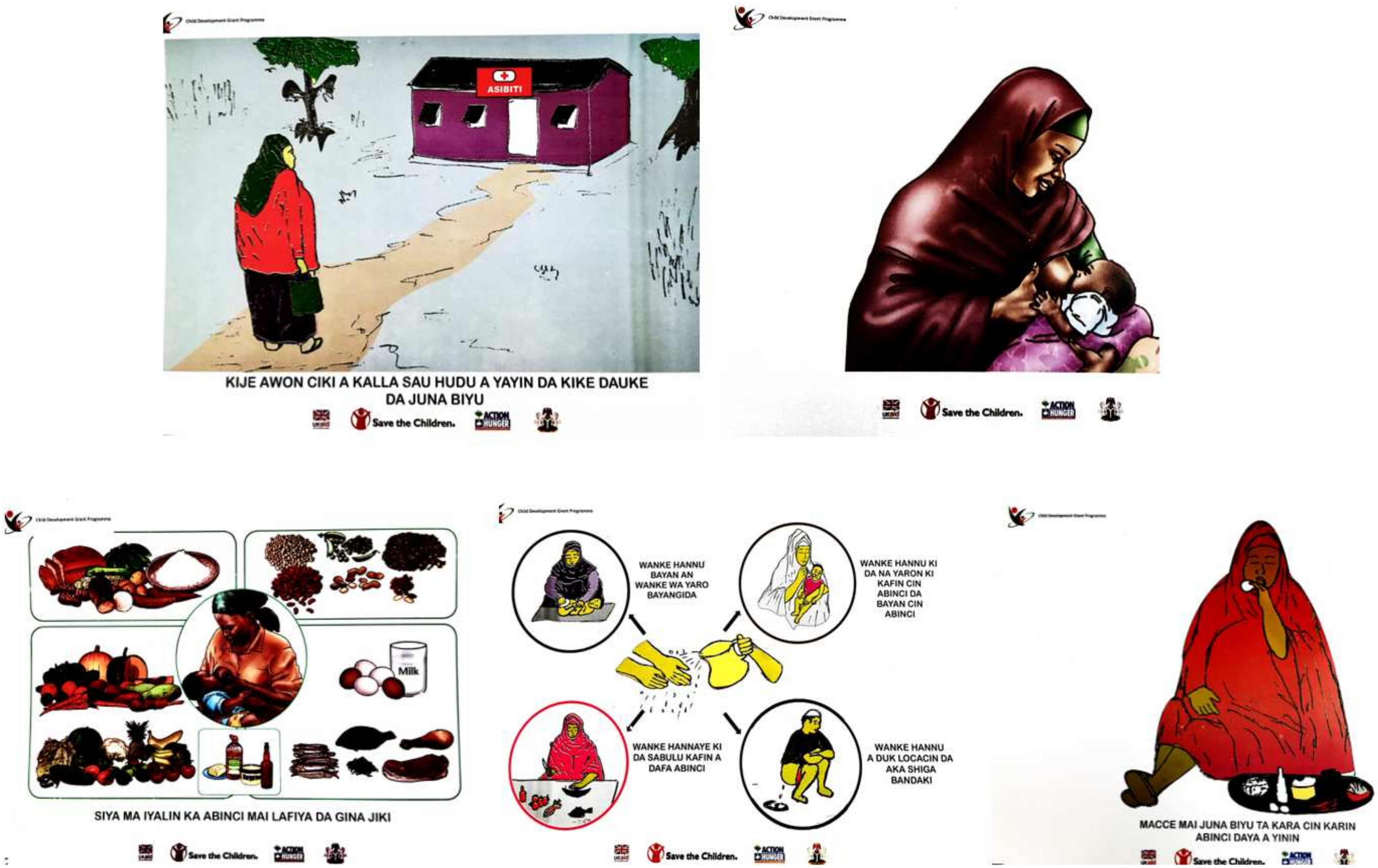
Notes: Significance levels: * (10%), ** (5%), ***(1%). The non-numbered first Column reports the ITT estimates at Midline for the new child, that have been reported in previous tables. Column 1 reports the mean (and standard deviation for continuous outcomes) for control households at Baseline for the Old Child. Column 2 reports ITT estimates at Midline for the old child. Column 3 reports the mean values in control households at Endline for the End Child. Column 4 reports ITT estimates at Endline for the end child. Each ITT is estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout.

Figure A1: Motivation



Notes: Panel A shows the number of individuals living in extreme poverty in 2018 (less than $1.90/day) [World Poverty Clock]. Panel B Shows the infant mortality rate per 1,000 [World Health Organization (WHO), Nigeria Demography and Health Survey 2013, and the World Bank]. Panel C shows the percentage of under fives's who are stunted (so their height-for-age-z-score (HAZ) is under -2 standard deviations of the WHO defined guidelines [WHO 2009]). The source of the data is the same as in Panel B. Panel D shows the diet diversity of children aged 0-5 in our data at Baseline, by decile of monthly food expenditure. The food groups are defined as: 1. Grains, roots and tubers, 2. Legumes and nuts, 3. Dairy products, 4. Flesh foods, 5. Eggs, 6. Vitamin-A rich fruits and vegetables, 7. Other fruits and vegetables.

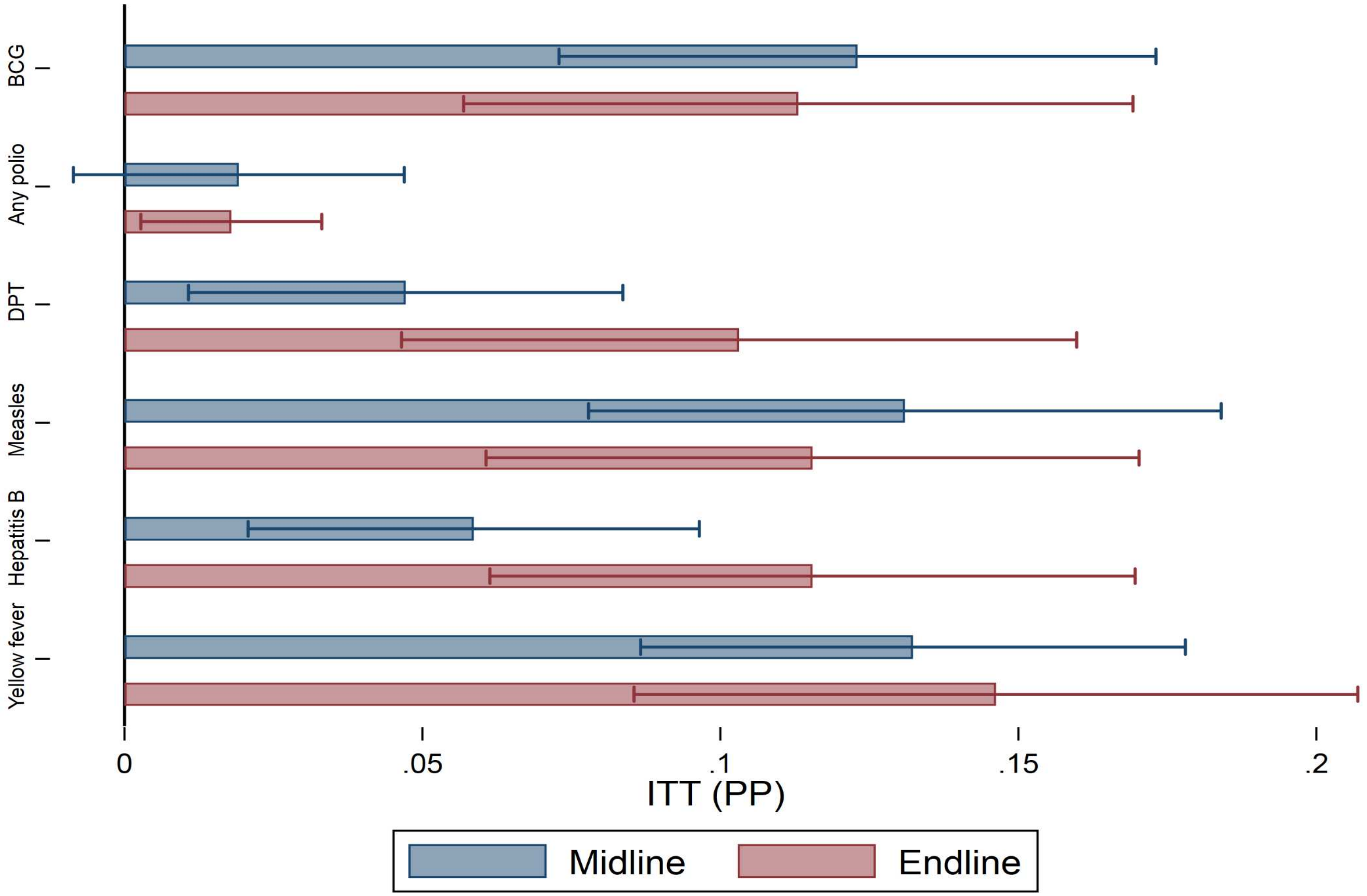
Figure A2: Examples of Visual Aid Materials



Note: Example of instructional materials from the program curriculum. Source: CDGP facilitator guide.

Figure A3: ITT Impacts on Vaccinations

Sample: Households with Pregnant Women at Baseline (N=3688)



Notes: Each bar shows the ITT estimates at Midline and Endline, along with the 95% confidence intervals on each. These are estimated using OLS, controlling for LGA and randomization tranche fixed effects, and the following Baseline characteristics of the household and mother: the number children aged 0-2, 3-5, 6-12 and 13-17, the number of adults, the number of adults aged over 60, mother's age, whether she ever attended school, total monthly expenditure, a dummy for polygamous relationships. Standard errors are clustered at the village level throughout.