United States Agency for International Development
Bureau of Democracy, Conflict and Humanitarian Assistance
Office of Food for Peace

Impact Evaluation of the Strengthen PSNP4 Institutions and Resilience (SPIR) Development Food Security Activity (DFSA)

Baseline Report

First Draft: March 26, 2019
Revised Draft: November 26, 2019

Harold Alderman, Fantu Bachewe, Daniel Gilligan, Melissa Hidrobo, Natasha Ledlie, Gayathri Ramani and Alemayehu Seyoum Taffesse

Submitted by
The International Food Policy Research Institute (IFPRI)*

On behalf of
IFPRI, Hawassa University, Ambo University,
Teachers College at Columbia University, Johns Hopkins University,
World Vision, ORDA, CARE

* We gratefully acknowledge funding for the impact evaluation by USAID under Cooperative Agreement No AID-FFP-A-16-00008. This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). Additional funding support for this study was also provided by PIM. Michael Mulford, Chief of Party of the SPIR project, made many contributions to the impact evaluation design and shares intellectual credit for overall research design. Helen Verdeli and Paul Bolton contributed to the design of the maternal depression substudy and its integration into the overall study design. For correspondence, contact Daniel Gilligan, International Food Policy Research Institute, 1201 Eye St., NW, Washington, DC 20005. email: d.gilligan@cgiar.org. phone: 202-862-8146.

Table of Contents

List of Tables

List of Figures

Acronyms

ANCOVA	Analysis of covariance
BCC	Behavior change communication
CAPI	Computer-Assisted Personal Interview
CF	Community facilitator
CPNP	Community Participatory Nutrition Promotion (CPNP)
CSA	Central Statistical Agency
DFSA	Development Food Security Activity
DA	Development agents
DID	Difference-in-difference
DS	Direct Support
EC	Ethiopian Calendar
FFP	Food for Peace
GOE	Government of Ethiopia
GMP	Growth Monitoring Promotion
GRAD	Graduation with Resilience to Achieve Sustainable Development
HAZ	height-for-age Z-score
HAD	Health development army
HEW	Health extension worker
HABP	Household Asset Building Program
IFPRI	International Food Policy Research Institute
IPT-G	Interpersonal Therapy in Groups
IYCF	Infant and young child feeding
MDE	Minimum detectable effect
MoLSA	Ministry of Labor and Social Affairs
MUAC	mid-upper arm circumference
OFSP	Other Food Security Program
ORDA	Organization for Rehabilitation and Development in Amhara
PCA	Principal components analysis
PDS	Permanent Direct Support
PIM	Program on Policies, Institutions, and Markets
PSNP	Productive Safety Nets Program
PSNP4	Fourth phase of the Productive Safety Nets Program
PW	Public Works
RCTs	randomized controlled trials
SAA	Social analysis and action
SNNP	Southern Nations, Nationalities, and Peoples

SPIR	Strengthen PSNP4 Institutions and Resilience
TDS	Temporary Direct Support
TTS	Timed and Targeted Counselling
VESA	Village Economic and Social Association
WASH	Water, sanitation and health
WHZ	weight-for-height Z-score

1. Introduction

Since its inception in 2005, the Productive Safety Net Program (PSNP) has been a cornerstone of the Ethiopian government's strategy for poverty alleviation, disaster risk management and rural development. The PSNP provides food or cash transfers targeted to poor households in the form of payments for seasonal labor on public works (PW) or as direct support (DS) to households whose primary income earners are elderly or disabled. With more than 8 million beneficiaries, the PSNP is one of the largest social protection programs in Sub-Saharan Africa. The PNSP has played an important role in improving the lives of poor Ethiopian households by reducing household food insecurity, increasing asset holdings and improving agricultural productivity (Berhane et al. 2014; Hoddinott et al. 2017). The fourth phase of the Productive Safety Net Program (PSNP4) began operating in 2016. Under PSNP4, the Government of Ethiopia (GOE) undertook a new round of targeting to identify client households. GOE also added a new objective to the PSNP to improve the nutritional status of women and children by better linking PSNP clients to health and nutrition services and through nutrition conditionalities (World Bank 2014). PNSP4 also included an enhanced livelihood transfer program to strengthen livelihoods and build assets, seeking to improve on the performance past livelihood transfer components of the PSNP.[1]

The Strengthen PSNP4 Institutions and Resilience (SPIR) Development Food Security Activity (DFSA) in Ethiopia is a five-year project (2016-2021) supporting implementation of the PSNP4 as well as complementary livelihood, nutrition, gender and natural resource management activities to strengthen the program and expand its impacts. Under funding from USAID's Food for Peace (FFP) Initiative and in close collaboration with the Government of Ethiopia, World Vision leads implementation of the SPIR DFSA, in partnership with the Organization for Rehabilitation and Development in Amhara (ORDA) and CARE. SPIR DFSA will target more than 500,000 PNSP clients in 15 of the most vulnerable woredas in Amhara, Oromia and SNNP regions of Ethiopia. SPIR DFSA also incorporates a substantial learning agenda component intended to use evidence to improve the design of the DFSA, provide feedback to strengthen its delivery and draw lessons both for local government and other national and international stakeholders about the potential to improve outcomes for PNSP clients through this type of expanded programming. World Vision, ORDA and CARE provide guidance on the implementation of the overall learning agenda. IFPRI leads the planning and execution of the learning agenda activities in collaboration with Hawassa University and Ambo University.[2]

The shared USAID Ethiopia, FFP, and PSNP4 objective for the DFSA is: resilience to shocks and livelihoods enhanced, and food security and nutrition improved, for rural households vulnerable to food insecurity. This objective is supported by DFSA activities under four purposes:

[1] Prior livelihood transfer components of the PSNP included the Other Food Security Program (OFSP) and the Household Asset Building Program (HABP).

[2] Bahir Dar University in Amhara also supports the SPIR Learning Agenda in an advisory role.

Purpose 1: Increased income, productive assets and equitable access to nutritious food for vulnerable women, men and youth;

Purpose 2: Improved nutritional status of children under two years of age, pregnant and lactating women, and adolescent girls;

Purpose 3: Increased women's and youth empowerment and gender equity;

Purpose 4: Strengthened ability of women, men and communities to mitigate, adapt to and recover from human-caused and natural shocks and stresses.

As part of the SPIR Learning Agenda,[3] IFPRI is collaborating with Hawassa University and Ambo University to design a structured, mixed methods impact evaluation to measure the causal impact of key activities of the SPIR DFSA program on livelihood, food security and nutrition outcomes. The mixed methods approach involves a quantitative experimental evaluation design to measure project impacts relative to a control group and qualitative assessments to inform participants' experience with the project, important contextual factors shaping impact and potential impact pathways. This Impact Evaluation Baseline Report focuses on the quantitative impact evaluation.

The specific learning questions that the impact evaluation will answer are the following:[4]

1. What is the impact of adding livelihood activities integrated with nutrition behavior change and WASH to PSNP on income growth, food security and graduation from the PSNP?
2. What is the impact of adding integrated livelihood, nutrition behavior change and WASH activities to the PSNP on child stunting prevalence?
3. What is the impact of combined livelihood and nutrition activities when both are integrated with activities to improve women's and youth empowerment to enhance women's impact on household diets and child nutrition relative to PSNP transfers alone?
4. Which combination of activities is most cost-effective for promoting graduation? And for reducing stunting?

The SPIR impact evaluation will provide valuable evidence for Ethiopia's policy makers in the Ministry of Labor and Social Affairs (MoLSA) and Ministry of Agriculture on how to expand the impact of PSNP4 through complementary interventions related to livelihoods and nutrition. In addition, the SPIR impact evaluation will contribute to the small but growing literature on the

[3] Implementation of the learning agenda follows USAID's methodology for collaborating, learning and adapting (CLA) to improve project effectiveness. The CLA approach extends traditional M&E practices and learning-based impact evaluations to develop a more integrated approach to communication between the project implementation, M&E and research teams, co-design of learning activities, and feedback to improve project delivery and effectiveness throughout the implementation period. See the SPIR DFSA Learning Agenda Inception Report (2017) for a more detailed description of how the SPIR Learning Agenda implements the CLA approach.

[4] For a more detailed list of learning questions, see the SPIR DFSA Impact Evaluation Inception Report (November 2017).

impact of graduation model social protection programs. Graduation model programs are comprehensive safety net programs that add a number of other activities related to asset building, income generation and access to markets in addition to traditional targeted cash or food transfers in order to provide a "big push" to help poor client households overcome potential poverty traps, hopefully leading to a permanent movement up the income distribution and graduation from the safety net. Graduation programs have gained prominence recently in part due to evidence from randomized controlled trials (RCTs) conducted in 6 countries (Ethiopia, Ghana, Honduras, India, Pakistan, Peru) showing that graduation programs built on the BRAC model led to large improvements in household economic outcomes, including consumption, food security, assets, finance and income (Banerjee et al. 2015). The Ethiopia SPIR graduation approach is distinct from these BRAC programs in two important ways. First, SPIR includes substantial, integrated programming designed to improve nutrition as well as women's and youth empowerment. The BRAC programs showed no effects on women's empowerment and nutrition outcomes were not assessed because this was not an objective of those programs. The impact evaluation of the SPIR project will be the first study we are aware of that will test the impact of a nutrition-sensitive graduation program. Second, most components of the SPIR graduation program operate through linking PSNP clients to financial services and markets and providing information and training as the primary strategy to improve livelihoods and other outcomes. These components of the program do not involve substantial resource transfers. The only exception to this is a counterpart to the PSNP4 livelihoods transfer, implemented on an experimental basis, in which the poorest half of client households in randomly assigned communities receive poultry packages or a large one-time cash grant. This stands in contrast to the BRAC models tested in Banerjee et al. (2015), in which the average value of resources transferred was 100% of household consumption, a doubling of income. Although the direct transfers through the SPIR project are smaller, the program activities are integrated and substantial, with great potential to provide new income generating opportunities and remove information constraints to improving wellbeing. The SPIR approach is intended to increase demand for local services and enhance their capacity, with the goal of making this approach more sustainable.

The purpose of the Baseline Report is to provide an overview of the impact evaluation and study design, summarize the data collected during the baseline survey to inform the context for the impact evaluation, and demonstrate that the baseline survey data are balanced across the experimental study design. This Baseline Report is organized as follows. Section 2 provides an overview of the SPIR DFSA project. Section 3 describes the impact evaluation design. Section 4 describes the baseline survey data collection. Section 5 summarizes the baseline survey data and tests for baseline balance. Section 6 concludes.

2. SPIR Intervention Description

World Vision, CARE and ORDA designed the SPIR DFSA to support delivery of PSNP4 while also developing and delivering multisectoral programming across the four project purposes in order to enhance livelihoods, increase resilience to shocks, and improve food security and nutrition for PSNP4 clients. The SPIR project will use community-level programming, training of government staff involved in public service delivery at the woreda (district) and kebele (subdistrict) level, and targeted livelihood transfers to support and strengthen PSNP4. Resource transfers received by SPIR participants will come primarily from transfers received from the PSNP4, as well as one-time livelihood transfers provided to the poorest PNSP clients to support livelihoods and promote graduation. Most other benefits of the SPIR project appear in the form of improved public service delivery and trainings to promote learning and support for community-level groups. For learning purposes, the SPIR impact evaluation combines major core components and innovative new activities under Purpose 1 on livelihoods and Purpose 2 on nutrition, along with selected activities under Purpose 3 on gender and youth and Purpose 4 on climate resilience, into a study design of overlapping interventions to learn what combination of activities has the greatest impact and is most cost-effective at improving SPIR outcomes.

The main SPIR livelihood activities under Purpose 1 include forming Village Economic and Social Associations (VESAs), financial literacy, agriculture and livestock value chain development, livelihood transfers in the form of a poultry start-up package or large cash grant, home gardening and forage production. SPIR uses VESA groups to encourage savings and improve access to credit. The VESA model was developed by CARE under the USAID Ethiopia Feed the Future-funded GRAD project. VESA groups include 25-30 members who are SPIR project participants, brought together as a foundation for all economic and social activities supported by the project. VESA groups include men and women (often the husband and wife from a single household). In addition to facilitating savings and lending, the SPIR project works with VESA groups to foster financial literacy, develop business skills, enhance production skills, improve social capital, and catalyze women's empowerment. SPIR staff work with VESA members engaged in livestock and crop value chains to conduct participatory market analyses for shoats (i.e., sheep and goats), poultry, staple crops, and high-value crops. VESA groups also serve as a platform for other trainings and services provided by Development agents, model farmers, and private sector actors. The SPIR project also believes that VESAs build social cohesion and create a safe and fertile environment for trainings on social and cultural norms.

The SPIR health and nutrition package under Purpose 2 includes integrated nutrition behavior change and communication (BCC) and water, sanitation and health (WASH) activities. BCC activities are organized under an intervention model referred to as Timed and Targeted Counseling (TTC). Under TTC, community Health Extension Workers (HEWs) and Development Agents (DAs) are trained to provide lessons in health posts at the community level and through household visits on topics including infant and young child feeding (IYCF) practices, and adolescent and maternal nutrition. Other topics include diversifying diets into sources of nutritious foods (including cooking demonstrations) and promotion of utilization of

health and nutrition services. The WASH component includes providing support to village-level WASH management activities, limited support to improving sanitation infrastructure (water sources and latrines) and implementation of the Community-led Total Sanitation and Hygiene (CLTSH) approach in which HEWs and DAs are trained to foster improvement community sanitation and hygiene and reductions in the practice of open defecation.

3. Evaluation Design

The impact evaluation will use a clustered randomized controlled trial (RCT) design to learn about 1) the effectiveness of the main activities related to Purpose 1 on livelihoods and Purpose 2 on nutrition; 2) the added benefit of enhancing the main livelihoods model with Social Analysis and Action (SAA) and aspiration activities; and 3) the added benefit of enhancing the Nutrition BCC and WASH activities with increased male engagement and interpersonal therapy interventions to reduce maternal depression.

3.1 Experimental Study Intervention Components

Before introducing the experimental design, we first explain the main livelihood and nutrition program activities and enhanced versions of these activities that make up the experimental intervention models:

Intervention L: SPIR livelihoods package: starting Village Economic and Social Associations (VESAs), financial literacy, agriculture and livestock value chain development, home gardening and forage production

Intervention L*: SPIR livelihoods package **plus** (i) Social Analysis and Action (SAA) to improve women's access to markets, (ii) aspirations promotion activities, and (iii) targeted poultry or cash livelihood transfers

Intervention N: SPIR health and nutrition package: training Health Extension Worker supervisors (HEWS) and other leaders on infant and young child feeding (IYCF) practices; nutrition social behavior change communication (SBCC); and water, sanitation and hygiene (WASH) practices

Intervention N*: SPIR health and nutrition package **plus** (i) intensive nutrition timed and targeted counseling (TTC); (ii) community-based participatory nutrition promotion (CPNP); (iii) male engagement in nutrition BCC and men's groups; and (iv) Interpersonal Therapy in Groups (IPT-G) interventions to reduce maternal (and paternal) depression

In this evaluation design, activities L and N represent the main SPIR project activities from Purpose 1 on livelihoods and Purpose 2 on nutrition, respectively. Activities L* and N* represent enhanced versions of these multifaceted activities designed to fill evidence gaps on the importance of gender equity, aspirations, large-scale livelihood grants, male engagement in nutrition BCC and strategies to address maternal and paternal depression. Below we summarize each of the additional components in L* and N*:

L* enhanced livelihood activities

- ***Social Analysis and Action (SAA):*** In the SPIR program, SAA will be used to enable individuals and communities to explore and challenge social norms, beliefs and practices around gender and nutrition that shape their lives. SAA is a community-led social change strategy that addresses constraints on women's role in intrahousehold decision making, mobility, and choice of livelihood activities, as well as restrictions on access to markets that derive from cultural and social norms. Examples of such constraints include limits on where women may travel or on their ability to sell goats in the market. The evaluation will test an integrated SAA model that seeks to unlock the potential of households to improve their wellbeing by providing women with greater voice, autonomy and access to economic activity.

- ***Aspirations:*** IFPRI researchers and others have conducted experiments in Ethiopia showing substantial and long-lived effects of an aspirations intervention based on documentaries designed to motivate individuals to undertake actions that will improve their wellbeing in the future.[5] These documentaries, in the Amharic and Afaan Oromo languages, will be used as an aspirations intervention within L*. The experimental design will randomize access to the aspirations intervention to households in half of the kebeles within the L* design, making it possible to separately identify the impact of the SAA approach within L*.

- **Enhanced livelihoods package for women:** Another component of SPIR Purpose 1 includes livelihood transfers in the form of poultry start-up packages or large one-time unconditional cash transfers. These transfers are targeted toward the poorest SPIR project participants, though not all of the poorest households will obtain these transfers. This project component is designed to mirror the PSNP4 targeted (rationed) livelihoods transfer, although targeting and programming of these transfers in the SPIR project differs from the PSNP4 approach. In the SPIR project, these livelihood transfers will be given to women in selected poor households either as a transfer of $200 in cash (ETB equivalent) or as $200 worth of poultry start-up inputs and training. Households receiving the cash transfer will be able to use it for any purpose without any instructions from the SPIR project. The poultry start-up package will include sasso breed pullets from

[5] See Bernard et al. (2017) and Taffesse and Tadesse (2017) for the results of recent aspirations experiments conducted in Ethiopia.

EthioChicken, feed, chicken coop construction materials, a feeding troth and training. The purpose of providing either cash grants or poultry start-up packages is to learn about which approach is more effective as improving livelihoods and other outcomes, as described below in the impact evaluation study design. These livelihood packages will be distributed in April 2019 after the aspirations documentaries have been shown in randomly selected L* kebeles (December 2018). This sequencing of interventions will allow us to test if receiving a positive aspirations shock prior to receiving a poultry business start-up kit or large cash grant changes use of the transfers and livelihood outcomes.

N* Enhanced Nutrition Activities

- ***Male engagement in nutrition BCC:*** Household level counseling (involving both husband and wife) related to IYCF and maternal nutrition will be conducted using the timed and targeted counseling (TTC) approach to support shared decision making. Because TTC conducts nutrition trainings directly in the community, it is more intensive than the SBCC provided in the SPIR nutrition package (N). SPIR will hire Community Health Facilitators (CHFs) for **each of the N* kebeles** to provide supportive supervision and monitoring of Health Development Army (HAD) volunteers in their household level counseling and other community health activities. These CHFs will also support the training of Community Participatory Nutrition Promotion (CPNP) for nutritious food preparation at Growth Monitoring Promotion (GMP) sessions using locally available and affordable foods to help rehabilitate underweight children. These trainings will also emphasize men's role in providing nutritious food at household level for children and mothers.

- ***Male engagement through men champions, men's groups, public awareness campaigns***: Male advocators will be identified and trained and will facilitate eight sessions for newly established men's groups in **each of the N* kebeles**. The sessions will provide an opportunity for men to critically reflect on cultural gender norms, gender relations, and explore the positive and perceived negative effects of male involvement, seeking to better understand how gender inequity affects the lives of women, children and men. Afterward, community level awareness events and public campaign on men engagement will be facilitated. These may involve video or drama presentations to increase engagement by men and boys in child care, household chores and even food preparation.

- ***Interpersonal Therapy in Groups (IPT-G) to address maternal and paternal depression:*** Recent studies have shown a strong association between maternal depression post-partum and child nutrition outcomes. One study recommended, "Interventions to promote growth in infants should include prevention or treatment of maternal depressive disorders and strategies to ensure adequate food security." (Wachs et al, 2009) Women who are screened to be suffering from maternal depression during the IFPRI-managed

midline survey (May-June 2019) will be invited to enroll in 12-week IPT-G sessions (approximately 5-8 women in each group). SPIR will hire IPT-G Officers for each woreda to supervise and support Community Health Facilitators to conduct these 12-week sessions in **each of the N* kebeles.**

3.2 Experimental Design

The impact evaluation design will compare combinations of these activity packages by randomly assigning kebeles to one of the following four intervention arms (see Figure 3.1.1):

- **Treatment 1:** L* + N*
- **Treatment 2:** L* + N
- **Treatment 3:** L + N*
- **Control:** PSNP only

Consistent with the graduation model design, the treatment arms in the experiment are integrated combinations of L, L*, N and N*. The evaluation will test the relative effectiveness of those combinations. We will find evidence on the impact of the fully integrated nutrition and livelihood models that include SAA, aspiration activities, livelihood transfers, male engagement in BCC, and IPT-G (T1) against the Control (answering a version of learning question Q1). We will also measure the impact of adding only male engagement and IPT-G (T2 v C) or adding only SAA, aspirations and livelihood activities (T3 v C) to the main SPIR model. These last two comparisons will tell us the additional effect of SAA, aspirations and livelihood transfers or of male engagement and IPT-G on top of the main integrated SPIR intervention. These comparisons will inform learning questions Q2 and Q3. Q4 will be answered by comparing the impact and cost of T1, T2 and T3.

Figure 3.1.1 SPIR Experimental Impact Evaluation

		Livelihoods	
		L*	**L**
Nutrition	**N***	T1 **L* + N***	T3 **L + N***
	N	T2 **L* + N**	T4 **PSNP Only**

3.2.1 Substudy on Cash Benchmarking of the Poultry Livelihood Intervention

The PSNP4 includes a Livelihood Transfer component, in which a fraction of the poorest PSNP beneficiaries, identified through community targeting, also receive an asset transfer designed to promote business development. The SPIR project will implement enhanced livelihood transfers in the form of a poultry start-up package or unconditional cash grant to mirror this feature of the PSNP4. These livelihood packages will be provided to the 10 poorest out of 18 households in each L* kebele in the SPIR study, with poorest households 10 selected according to an asset index developed from the baseline data.[6]

The poultry start-up package was selected in part because of the availability of the promising Ethiochicken breeds, which appear to be highly productive. In addition, there is renewed international attention on poultry as an asset which is widely accessible to women and which has low start-up costs. In 2106, Bill Gates promoted investment in chickens to help increase incomes for poor women (https://www.gatesnotes.com/Development/Why-I-Would-Raise-Chickens). In response, Chris Blattman suggested that large cash grants of the kind provided by Give Directly (Haushofer and Shapiro 2016) may be effective at improving outcomes for more women given heterogeneity in their needs and capacity to raise chickens (https://www.vox.com/the-big-idea/2017/3/14/14914996/bill-gates-chickens-cash-africa-poor-development). This debate sparked interest in testing promising development interventions like poultry start-up packages by benchmarking them against cash transfers of similar value.[7] The data from the SPIR impact evaluation will provide an opportunity to contribute evidence on this debate, comparing valuable

[6] The asset index was constructed using ownership data on more than 30 asset categories including consumer durables, productive assets, livestock and land. The asset index was constructed using principal components analysis, which reduces the influence of ownership of assets in the index that are shown to be highly correlated with ownership of other assets (Filmer and Pritchett 2009).

[7] See McIntosh and Zeitlin (2018), for example, who conducted an experiment to benchmark a nutrition intervention in Rwanda against large cash grants.

Ethiochicken poultry start-up packages to an equivalent cash grant, within the context of an integrated graduation model social protection program. Although several other interventions will be taking place in the SPIR study, the randomized assignment of poultry or cash grant packages to the poorest households in the study will make it feasible to identify the impacts of either livelihood package, when combined with related complementary interventions that are also experimentally assigned to poultry and cash grant recipients.

3.2.2 Substudy on Maternal Depression

Recent evidence has identified maternal depression, particularly in the post-partum period, as a potentially important determinant of child growth and development outcomes in low income settings. In one study, infants of mothers with depressive symptoms had 2.17 higher odds of being stunted (95% CI: 1.24, 3.81) than did infants of mothers with few symptoms (Wachs et al. 2009). A prior study showed that lowering depression can reduce child stunting by 27 percent (Black et al. 2009). It has also been shown that an interpersonal therapy in groups (IPT-G) intervention was highly effective at reducing depression in Uganda (Bass et al. 2006; Bolton et al. 2003), and that an IPT-G intervention significantly reduced depression for adolescent girls (but not adolescent boys) living in internally displaced persons camps in war-affected northern Uganda (Bolton et al. 2007).

Addressing maternal depression is consistent with the SPIR project objectives under Purpose 3 to strengthen the capacity of women to improve outcomes for themselves and their families. Based on this evidence, the implementation and study teams agreed to implement a substudy on the impact of IPT-G on maternal depression and child nutrition outcomes. Psychologists Lena Verdeli (Teachers College, Columbia University) and Paul Bolton (Johns Hopkins University) have joined the research team as co-principal investigators on aspects of the study related to maternal depression.

During the midline, the mother of one child in each household that is age 0-35 months and her male partner, if any, will be screened for depression symptoms and functional effects of depression using the PHQ-9 symptom assessment tool. The PHQ-9 asks subjects to report the frequency with which they experienced each symptom of depression (e.g., feeling bad about yourself; feeling that you would be better off dead) over the previous two weeks, with coded responses ranging from 0 (not at all) to 3 (nearly every day). The tool yields a depression severity score from 0-27, with severity classified by intervals of: none (0), minimal (1-4), mild (5-9), moderate (10-14), moderately severe (15-19), and severe (20-27). In order to learn about the prevalence of depression in the study area and to gain experience with the PHQ-9 tool, one woman and one man from each household in the baseline survey was screened for depression using a slightly modified version of the PHQ-9. Results are reported below. Dr. Verdeli will lead training of specially recruited mental health workers (in January 2019) who will train community facilitators to lead IPT-G groups. After women and men are screened for depression in the midline survey in May-June 2019, those who are identified as depressed will be invited to participate in a 12-week IPT-G session. The study will assess whether this approach to treating depression is effective, whether alleviating depressive symptoms has benefits for child nutrition

and health status. In addition, we will assess whether the SPIR treatment arms contributed to reduced depression and whether lower depression is associated with larger impacts of the SPIR interventions.

3.3 Implementing the SPIR evaluation study design

The experimental study design involves randomly assigning 193 kebeles into these four treatment arms (see Section 3.4 for details). All kebeles assigned to L* (T1 and T2) (n=96) will receive the SAA intervention. In addition, study households in half of the L* kebeles (n=48) will receive the aspirations intervention, stratified across T1 and T2. The livelihood transfer of the poultry start-up package or unconditional cash grant will be provided to the 10 poorest households in the L* study communities. Randomization of poultry or cash grants will be done at the kebele level, with women in one half of the L* kebeles randomly assigned to receive the poultry-start up package, and women in the remaining half of the L* kebeles (48) to receive the cash grant. Randomization of the poultry/cash livelihood intervention will be balanced across the aspirations and non-aspirations kebeles in the L* intervention arms.

All households in kebeles assigned to N* (T1 and T3) will participate in the male engagement intervention. These households will receive both targeted male engagement through timed and targeted counseling (TTC) household visits that specifically include men, together with their spouses, the CPNP promotion, and the men's champions/groups and public awareness campaigns. In addition, women who are mothers of children age 0-35 months and their male partners will be screened for depressive symptoms during the midline survey and invited to enroll in IPT-G if found to be depressed, as described in Section 3.2.2.

3.4 Timeline of activities

While the L and N activities were rolled out soon after the baseline survey in 2018, the L* and N* activities are planned to be rolled out in late 2018 and in 2019. SAA will be rolled out in all L* kebeles, the aspirations intervention in half of the L* kebeles, nutrition BCC with a men's engagement focus and IPT-G in all N* kebeles in 2019 (see Figure 3.4.1).

The quantitative impact evaluation will include three rounds of household data collection, a baseline survey, a midline survey and an endline survey. The baseline survey was conducted from January 25 – April 9, 2018, before SPIR activities were rolled out in study areas.[8] The midline survey will be conducted in May-June 2019, one year after the baseline and the endline survey will be conducted two years after the baseline, in May-June 2020.[9] The midline and endline surveys will collect data on the same households from baseline, thus forming a 3-year panel survey. In addition, the midline survey will include a supplemental sample of households in which a woman has given birth in the last 6 months in order to inform the maternal depression

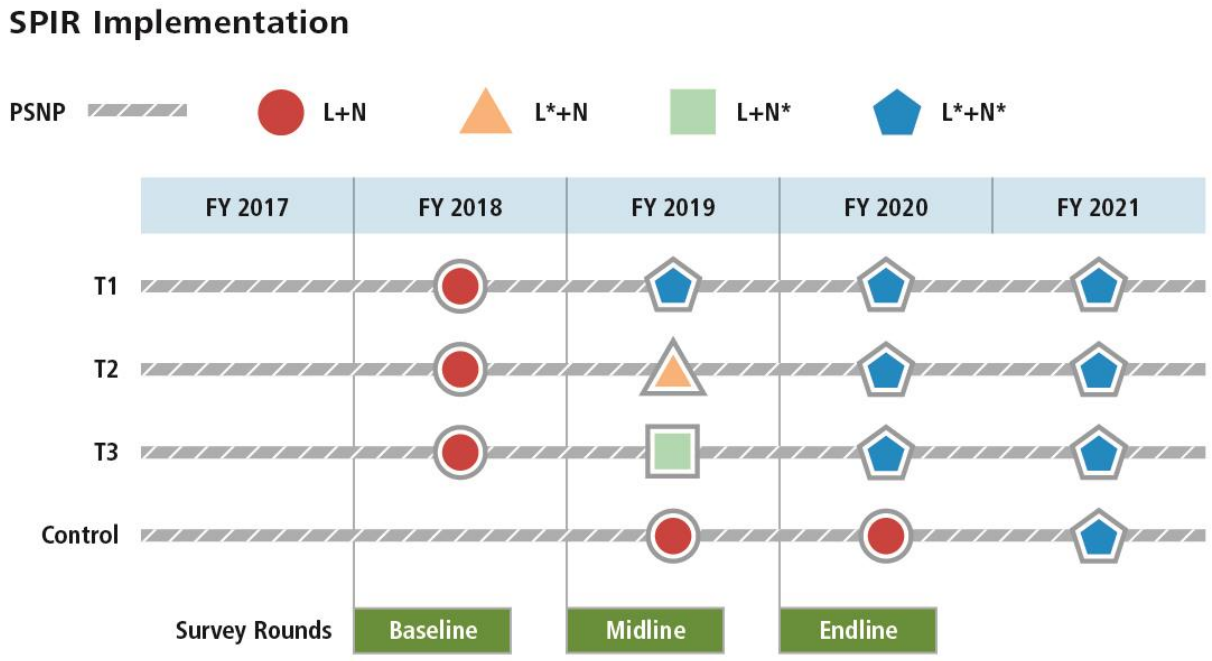
[8] SPIR project activities, particularly the formation of VESA groups, had already begun in many communities before the baseline survey, but these communities were omitted from the study.

[9] The midline survey was originally planned for March-April 2019 to overlap seasonally with the timing of the baseline survey, but the midline survey had to be delayed to accommodate the national census which is planned in April 2019.

intervention in N* and to be sure to have enough women enrolled in the IPT-G intervention. In this sense, the midline becomes a baseline survey for the part of the sample including these new mothers. The IPT-G interventions are only likely to be effective on adults displaying at least moderate signs of depression. Thus, implementing the IPT-G component would require screening new mothers to identify those suffering from post-partum depression. Screening of mothers would occur during the midline data collection, before the IPT-G component is rolled out.

This community-based approach to addressing maternal depression will require some local piloting and adaptation of tools, which is planned for early 2019. Also, psychologist Lena Verdeli (Teachers College, Columbia University), and a co-trainer will train IPT-G officers and other SPIR staff in a practical training, where these officers will be supported and supervised in conducting a test 12-week session from March to May. In May, these IPT-G officers will receive a five-day training in supervision and monitoring skills, as they will be the staff directly support the Community Health Facilitators who will be facilitating the sessions at the kebele level. After the initial 12-week session for those mothers screened during the May 2019 midline, there will be a second screening and enrollment later in the year.

Figure 3.4.1 Timeline of activities



3.5 Household Eligibility and Sampling

3.5.1 Kebele and Household eligibility criteria

The study takes place in 13 woredas across Amhara and Oromia regions of Ethiopia, where implementing partners – World Vision, CARE, and ORDA – planned to work. In designing the study sample, we began with a list of all kebeles in which the PSNP operated in these woredas. From each woreda, we dropped kebeles where Village Economic and Saving Association (VESA) program had already been formed, the first step in implementation of the SPIR project. A total of 196 kebeles (115 in Amhara and 81 in Oromia) remained as a part of the study. Two of the 196 kebeles were subsequently dropped for having no PSNP clients, and one kebele (Ejartii in Daro Lebu) was later dropped for security reasons.

In each kebele, we randomly sampled 18 households, leading to a planned baseline sample of 3,474 households. The inclusion criteria for the sample was that households had to (1) be a PSNP client household, (2) have at least one child age 0-35 months, and (3) have the mother or primary female caregiver of the 0-35-month-old child be a member of the household. The last criterion ensures that we can measure the relationship between mother and child diets and between maternal depression and child outcomes in all sample households. To know which households met the sampling criteria, we used World Vision's Master Beneficiary List of PSNP beneficiaries as the sampling frame and a validation exercise was conducted with each sampled household to confirm that the met the sample inclusion criteria before starting the household survey. If a household was found ineligible at the start of the interview, the enumerator replaced the household in the sample with the next household sampled from the Master Beneficiary List.

3.5.2 Sample size calculations

Power calculations were conducted to estimate the necessary sample size required to measure a detectable effect of the program on three outcomes of interest: child height-for-age Z-scores (HAZ), mother's nutrition knowledge, and food security as measured by the household food gap in months. Child HAZ is known to both indicate undernutrition and predict health and well-being in later years (Black et al. 2013). We obtained means, standard deviations (SDs), and intracluster correlations[10] from the PSNP4 midline round of data collection. The sample size estimation was based on the kebele-level randomization leading to comparisons of each treatment group with the control group.

We conducted the power calculations for ANCOVA models, adjusting for autocorrelation from baseline to endline. For HAZ, we assumed an autocorrelation of 0.7. For nutrition knowledge and a household's food gap we assumed no autocorrelation. We used conventional levels and set the power at 80 percent and the significance level at 0.05. Table 3.5.1 shows the minimum detectable effect size and full set of parameters as assumed or calculated while conducting

[10] The intracluster correlation is the fraction of the total variance of an outcome that can be explained by the within cluster variance.

sample size calculations. For HAZ, the minimum detectable effect (MDE) size is 0.34 SD; for mother's nutrition knowledge it is roughly 1.0 point on a mean score of 7.51; and for the food gap, it is 0.89 months from a mean of 2.35 months.

Table 3.2.1: Sample size calculations

Parameter	Child HAZ (under 24 months)	Mother's nutrition knowledge	Household food gap
Power	0.80	0.80	0.80
Size (alpha)	0.05	0.05	0.05
Number of clusters	196	196	196
Observations per cluster	18	18	18
Total baseline observations	3528	3528	3528
Minimum detectable effect size	0.34	0.99	0.89
Mean control	-1.73	7.51	2.35
Standard deviation control	2.27	3.42	2.54
Intra-cluster correlation	0.08	0.21	0.34
Autocorrelation	0.7	0	0

3.6 Randomized Assignment

3.6.1 First stage randomization

Of the 196 study kebeles that were randomized, 49 were assigned to each of the four treatment arms: T1: L*+N*; T2: L*+N; T3: L+N*; and C: PSNP only. Randomization was stratified at the woreda level to provide balance of treatment assignment geographically and because the woreda is the main administrative structure for local government which shapes local public expenditure and public service delivery. The first version of the randomization was conducted based on kebele-level implementation data provided in September 2017. Kebeles in which VESA groups had already formed were removed from the sample, leading to a total of 158 clusters. In November 2017, we were given a second dataset that had an additional woreda in Oromia, Daro Lebu, where implementation would start a few months later, which added 37 additional clusters to the study. In addition, original kebele level data on VESA group formation was incorrect in 4 of the Oromia kebeles. As a result, it was agreed to redo the randomization for Oromia region only, since VESA program formation had already begun in Amhara based on the initial randomization. In the second dataset, VESA programs in Oromia that were marked as "very new" were now kept in the eligible list of kebeles. Thus, the original randomization for the 115 kebeles in Amhara was kept and the new 81 kebeles in Oromia were re-randomized, ending up with a total of 196 clusters.

Using the initial data, 1,000 potential treatment allocations were generated, stratified by woreda. A uniform random number between 0 and 1 was drawn for each of the clusters in each stratum. These were ordered and then allocated 1/4th of the sample to each treatment arm. Since some strata did not consist of clusters that were evenly divisible by four, we randomly allocated the leftover clusters within each stratum to one of the treatment arms ensuring that balance across arms within strata would be preserved (i.e., no treatment arms gets more than one leftover cluster

within the strata) and that the allocation would be random. When there was one leftover cluster, a random number between 0 and 1 was drawn, and if it was less than 0.25 it was allocated to treatment group 1, if it was between 0.25 and 0.5 it was allocated to treatment group 2, if between 0.5 and 0.75 it was allocated to treatment group 3, and if between 0.75 and 1 it was allocated to treatment group 4. With two leftover clusters, there were 6 possible allocations across the 4 treatment groups (4 choose 2). Again, we drew a random number between 0 and 1 at the strata level, and if this was less than 0.1667 then the clusters go in treatment groups 1 and 2, if between 0.1667 and 0.3333 then the clusters go in treatment groups 1 and 3, and so on. A second random number is selected and ranked to decide the order of the allocation to each of the treatment arms. The same procedure was followed for strata with 3 leftover clusters.

Using the share of PSNP beneficiaries in each kebele and the distance from the kebele to the district capital to balance the treatment arms, the relative efficiency of each of the 1,000 potential allocations was calculated. For any treatment allocation, the relative efficiency provides a measure of the balance in observable characteristics between potential treatment groups. The maximum t-statistic from the regression of the observed characteristic on the treatment allocations (with strata dummies) is calculated at the sample level. Allocations with the most equal allocations across regions were kept from these 1,000 allocations – that is, allocations with more than 29 kebeles per treatment arm in Amhara were dropped, and allocations with more than 12 kebeles per treatment arm in Oromia were dropped. At the sample level, allocations that resulted in less than 39 clusters in each treatment arm were also dropped. From the remaining allocations, the one with the highest relative efficiency – the minimum maximum t-statistic - was retained (Bruhn and McKenzie 2009). This allocation is used as the final randomization allocation for Amhara.

In the second set of data, that we received on November 1, the procedure was modified to take as given the previous assignment of kebeles to treatment groups in Amhara. For each stratum in Oromia, 1,000 potential treatment allocations were generated using the same procedure that was used in the initial randomization. Leftover clusters within each stratum were also dealt with in the same way as the initial randomization. For each of the 1,000 potential treatment allocations generated for Oromia in the second set of randomizations, the relative efficiency is calculated using the potential treatment allocation for previously unassigned kebeles in Oromia but the actual treatment assignment for kebeles in Amhara that were allocated to a treatment groups during the previous wave of the randomization. Allocations with the most equal balance across Oromia were kept – that is allocations that resulted in a treatment group with 19 or fewer clusters were dropped. At the sample level, allocations that resulted in 49 clusters per treatment arm were kept. From the remaining allocations, the one with the highest relative efficiency – the minimum maximum t-statistic - was retained. This allocation is used as the final randomization allocation, maintaining the original Amhara randomization and combining it with the new randomization for Oromia (Table 3.3.1).

Table 3.3.1: Number of kebeles in each treatment arm, by woreda

Region	Woreda	T1: L*+N*	T2: L*+N	T3: L+N*	C: PSNP4	Total
Amhara	Bugna	2	1	1	1	5
	Dahena	3	4	4	4	15
	Gaz Gibla	2	1	2	2	7
	Lasta	3	4	4	4	15
	Meket	10	9	9	10	38
	Sekota	5	5	5	4	19
	Wadla	4	4	4	4	16
Oromia	Chiro	3	3	3	3	12
	Daro Lebu	9	10	9	9	37
	Gemechis	2	3	3	3	11
	Grawa	4	3	3	4	14
	Kurfachelle	0	1	1	0	2
	Siraro	2	1	1	1	5
Total		49	49	49	49	196

Note: After the sampling was completed, two new woredas – Gazo and Tsagabji – were created from the existing woredas. This led to a reshuffling of some kebeles in Meket, Wadla and Lasta. For the purpose of this report and for subsequent analysis, we refer to these kebeles by the woreda to which they belonged at the time of sampling. Kebele treatment assignments and associated implementation remain unchanged despite these administrative changes.

After the initial randomization of kebeles across the four treatment arms, two kebeles were dropped because they had no PSNP beneficiaries and one was dropped for security reasons. The randomization across the 193 remaining kebeles is shown in Table 3.3.2.

Table 3.3.2: Number of kebeles in each treatment arm, by region

	Amhara	*Oromia*	*Total*
T1: L*+N*	27	19	46
T2: L*+N	28	21	49
T3: L+N*	29	20	49
C: L+N	29	20	49
Total	113	80	193

3.6.2 Second stage randomization

Within the 95 L* kebeles assigned to the treatment arms T1: L* + N* and T2: L*+N, half were randomized, by woreda, to (1) receive aspirations treatment; and (2) not receive aspirations treatment. Again, 1000 potential treatment allocations were calculated, by woreda, of which only allocations with balance between the number of kebeles in the aspirations treatment in L*+N* and L*+N treatment groups, were kept. Of the allocations that remained, a random number between 0 and 1 was assigned to each randomization and the one with the lowest random number assigned was kept as the realized allocation.

The L* kebeles were also randomized into poultry and cash treatment arms. Like the aspirations treatment assignment, 1000 potential allocations were drawn, stratified by woreda. Only the allocations with balance between the number of poultry treatment kebeles in L*+N* and L*+N treatment groups, as well as balance between the number of poultry treatment kebeles in L*+N* aspirations vs non-aspirations, and L*+N aspirations vs non-aspirations groups were kept. Of the allocations that remained, a random number between 0 and 1 was assigned to each randomization and the one with the lowest random number assigned was kept as the realized allocation.

The resulting final kebele-level randomization is presented in Tables 3.3.3 and 3.3.4.

Table 3.3.3: Kebele level sub-randomizations in T1: L*+N* treatment arm

	Cash	Poultry	Total
No aspirations treatment	12	11	23
Aspirations treatment	11	12	23
Total	23	23	46

Table 3.3.4: Kebele level sub-randomizations in T2: L*+N treatment arm

	Cash	Poultry	Total
No aspirations treatment	12	13	25
Aspirations treatment	12	12	24
Total	24	25	49

3.7 Targeting of households within the poultry and cash intervention

The poultry and cash intervention is targeted to the ten poorest households in each of the 95 L* kebeles based on a wealth (asset) index constructed from baseline data. The household wealth index derived from different measures of asset ownership using Stata's *pca* command for principal component analysis of the correlation matrix. We combined consumer durable assets, productive assets, livestock assets, and size of land owned, scores from which the first component explained 11.3% of the variance. Using the scores of this wealth index, we ranked all households within their kebele, to pick the poorest ten households in each kebele. The overlap is high between the ten poorest households selected using this wealth index and indices created using asset indices for the individual asset categories, as well as comparing it to the monthly consumption expenditure per adult equivalent, indicating that the targeting of the poorest ten households is only modestly sensitive to the index that is chosen.

3.8 Empirical strategy

Impacts of the SPIR project will be estimated on the baseline and endline data using Analysis of Covariance (ANCOVA) models, with single difference models and difference-in-difference (DID) models (for outcomes with baseline imbalance) used for robustness checks. The ANCOVA model is more flexible than a DID model because the ANCOVA model allows the

autocorrelation in the outcome over time to be estimated, rather than fixed at one, as in the DID model. This provides a better model fit (McKenzie 2012). Moreover, there are statistical power gains from using ANCOVA models over DID models which get larger as the autocorrelation in the outcome falls. When autocorrelation in the outcome is low, which is likely for many of the outcome variables in this evaluation, the benefit in statistical power from using ANCOVA is substantial.[11]

Using the ANCOVA model, we will estimate intent to treat (ITT) effects of the SPIR program by estimating the average impact of SPIR on a random sample of beneficiaries, regardless of whether they participate in all aspects of the intervention for their relevant treatment arm. Study subjects in all four treatment arms will be PSNP4 beneficiaries, as even the households selected for the study in the Control kebeles are in the PSNP4, but access to the other SPIR components in each of the other treatment arms could vary within a kebele. The ITT effect captures differences in coverage of the program within communities or decisions by beneficiaries not to participate in all aspects of the program. The average treatment effect (ATE), on the other hand, is the actual effect of the full intervention for that treatment arm on households that receive it. However, because compliance is not perfect, and not all beneficiaries receive all components of the program for their treatment arm, using the ITT effect is the better approach.

To maximize statistical power, we can estimate the impact of the **combined treatment group** $T \in \{T1, T2, T3\}$ against the Control group, using the following empirical specification for the ANCOVA model:

$$Y_{1hv} = \beta_0 + \beta_1 T_v + \beta_2 Y_{0hv} + \beta_3 X_{0hv} + \varepsilon_{hv}, \qquad (1)$$

where Y_{1hv} is the outcome of interest at endline for household *h* from kebele *v*, Y_{0hv} is the outcome of interest at baseline, and T_v is an indicator for whether kebele *v* was assigned to any of the SPIR treatment arms and X_{0hv} is a set of control variables at baseline for household *h* from kebele *v*. β_1 measures the impact of the combined treatment as the difference in the average outcome between the treatment arms T1, T2 and T3 combined and the Control group.

To estimate the impact of each treatment arm separately against the Control group using the ANCOVA model, we estimate the following model:

$$Y_{1hv} = \beta_0 + \beta_1 T1_v + \beta_2 T2_v + \beta_3 T3_v + \beta_4 X_{0hv} + \varepsilon_{hv,} \qquad (2)$$

[11] The ratio of the difference in differences variance to the ANCOVA variance is 2/(1+ρ), where ρ is the autocorrelation. When ρ=.25, with a single baseline survey and follow-up survey, the sample size needed is 60 per cent larger with a DID model than with ANCOVA to get the same power.

where $T1_v$ is an indicator for whether the household in kebele *v* was randomly assigned to treatment T1, $T2_v$ indicates randomized assignment to T2 and $T3_v$ indicates randomized assignment to T3. β_1, β_2, and β_3 measure the impact of T1, T2 and T3 respectively. To test whether the ITT estimators are statistically different across treatment arms T1 and T2, for example, we conduct a Wald test of equality of the estimates β_1 and β_2.

The evaluation will also differentiate household impacts of the most intensive treatment, T1, for example, by whether the household was randomly assigned to the aspirations documentary treatment or not. This impact will be estimated on the sample of households in T1 or the Control group only using the following model:

$$Y_{1hv} = \beta_0 + \beta_1 T1_v * A_{hv} + \beta_2 T1_v + \beta_3 Y_{0hv} + \beta_4 X_{0hv} + \varepsilon_{hv}, \qquad (3)$$

where $T1_v * A_{hv}$ is an indicator for household *h* from kebele *v*, being assigned to the aspirations treatment within a kebele assigned to T1, and $T1_v$ indicates a household not assigned to the aspirations treatment in a kebele assigned to T1. We can test whether the aspirations treatment increases the impact of T1 relative to the control group by testing equality of β_1 and β_2.

The absolute and relative impacts of the SPIR project may depend on baseline characteristics of the study sample. Amhara and Oromia regions differ in terms of agriculture, nutrition, gender norms, natural resources and markets. Consequently, we plan to measure heterogeneity of impact by region, following Bruhn and McKenzie (2009).

4. Baseline Data Collection

Fieldwork for the baseline survey data collection was mostly completed from February 8 – April 25, with a small number of additional interviews and callbacks completed in the ensuing weeks. BDS Center for Development Research (BDS-CDR) served as the in-country survey partner, leading the baseline data collection in cooperation with the quantitative evaluation team from IFPRI.

4.1 Survey instrument

The baseline household questionnaire was designed by the IFPRI team based on substantial past experience conducting quantitative evaluations of PSNP and other agriculture and nutrition interventions in Ethiopia. The baseline household questionnaire was structured in three parts: a brief household-level interview for identification and household demographics, a male respondent questionnaire and a female respondent questionnaire. The baseline survey collected information on primary and secondary outcomes, basic demographics, indicators that were likely

to be predictive of the primary and secondary outcomes, and intermediate outcomes that are relevant for testing different causal mechanisms. GPS coordinates were also collected for each household to assist in tracking households in the midline and endline surveys. The list of baseline survey questionnaire modules is presented in Tables 4.1.1 – 4.1.3. The complete baseline questionnaire can be found in Appendix A.

The baseline household interview took approximately two hours to complete and required the mother or primary female caregiver of the index child ("primary female") and her husband or partner ("primary male") of each sampled household to respond to different questionnaire modules. Primary female and male respondents were selected based on the selection of the index child. In households with a single adult female and no adult male, some of the modules for the primary male respondent were skipped. Households with no adult female were not eligible for inclusion in the study. In households with both a target male and target female respondent, select modules were repeated for intrahousehold comparison.

In the final module of the survey, anthropometry measurements were collected by trained anthropometry specialists and the index child, sibling of the index child, and primary female were measured. Children that were found to have a low mid-upper-arm circumference indicating possible severe acute malnutrition were referred to the local health post in the kebelle.

The baseline household survey questionnaire was administered by enumerators using tablets with a computer assisted personal interview (CAPI) programmed in CSPro. The CAPI enabled enumerators to easily access pre-loaded data, follow interview skip patterns according to interviewee responses, and back-up survey data after each day of interviews.

Table 4.1.1: Baseline Survey Household Questionnaire Modules

Strengthen PSNP4 Institutions and Resilience (SPIR)
Development Food Security Activity (DFSA)
BASELINE SURVEY: Household Questionnaire – February 7, 2018

DRAFT: For Research Purpose only

Outline:

Module A: Household Identification and consent
Part 1: Household identification, location, and consent
Part 2: Sample verification

Module B: Household Composition and Characteristics
Part 1: Household Roster
Part 2: Child Schooling (age 6-18)

Module C: Result of Randomization
Part 1: Result of Randomization

Universal Codes (Include with all CAPI options):

97=Refuse to respond	98=Don't know	99= Not applicable

Table 4.1.2: Baseline Survey Male Questionnaire Modules

Strengthen PSNP4 Institutions and Resilience (SPIR)
Development Food Security Activity (DFSA)
BASELINE SURVEY: Male Questionnaire – February 7, 2018

DRAFT: For Research Purpose only

Outline:

Module A: Household Assets
- Part 1: Productive Assets
- Part 2: Consumer Durables

Module B: Livestock Production
- Part 1: Livestock
- Part 2: Income from Livestock and Specified Agricultural Products
- Part 3: Cost of Livestock Production
- Part 4: Agricultural Extension Related to Livestock

Module C: Agriculture
- Part 1: Land characteristics and tenure
- Part 2a: Crop choice – Belg Season
- Part 2b: Crop inputs and labour – Belg Season
- Part 2c: Crop production, sales, and use – Belg Season
- Part 3a: Crop choice – Mehr Season
- Part 3b: Crop inputs and labour – Mehr Season
- Part 3c: Crop production, sales, and use – Mehr Season
- Part 4: Agricultural Extension Related to Crop Production

Module D: Income Apart from Own-Agricultural Activities: Wage Employment

Module E: Own Business Activities
- Part 1: Own Business Activities
- Part 2: Training

Module F: Sources of Information

Module G: Access to credit and financial services
- Part 1: Credit for production purposes
- Part 2: Credit for consumption purposes
- Part 3: Access to savings
- Part 4: Access to insurance

Module H: Non-food expenditure
- Part 1: Durables and services (annual)
- Part 2: Household consumables (monthly)

Module I: Household decision-making and empowerment
- Part 1: Role in Decision-making
- Part 2: Agency
- Part 3: Intrahousehold Dynamics and Attitudes

Module J: Nutrition, Health, and Care of Child
- Part 1: IYCF Knowledge
- Part 2: Child care activities
- Part 3: Exposure to health and nutrition services

Module L: Aspirations and Wellbeing
- Part 1: Aspirations

Module K: Access to the PSNP and SPIR activities
- Part 1: Public Works
- Part 2: Direct Support
- Part 3: Understanding of PSNP4 operations
- Part 4: Other pubic transfers
- Part 5: Livelihoods component of PSNP4
- Part 6: Participation in VESA groups and SPIR activities

Module L: Aspirations and Wellbeing
- Part 2: Experience with shocks
- Part 3: Poverty perceptions and wellbeing
- **Part 4: Experience with Depression and emotional wellbeing**

Universal Codes (Include with all CAPI options):

97=Refuse to respond	98=Don't know	99= Not applicable

Table 4.1.3: Baseline Survey Female Questionnaire Modules

Strengthen PSNP4 Institutions and Resilience (SPIR)
Development Food Security Activity (DFSA)
BASELINE SURVEY: Female Questionnaire – February 7, 2018

DRAFT: For Research Purpose only

Module A: Housing, Water, Sanitation and Hygiene
- Part 1: Housing and water
- Part 2: Sanitation and hygiene

Module B: Assets Owned by the Woman
- Part 1: Consumer durables owned by the woman

Module C: Livestock Owned by the Woman
- Part 1: Livestock owned by the woman
- Part 2: Cost of livestock production
- Part 3: Income from livestock and specified agricultural products
- Part 4: Agricultural extension related to livestock

Module D: Own business activities

Module E: Sources of Information

Module F: Access to credit and financial services
- Part 1: Credit for productive purposes
- Part 2: Credit for consumption purposes
- Part 3: Access to savings

Module G: Non-food expenditure
- Part 1: Household consumables (monthly)

Module H: Food consumption and expenditure
- Part 1: Women's dietary diversity (24-hour recall)
- Part 2: Household food consumption and expenditure
- Part 3: Household food security

Module I: Household activities, decision-making and empowerment
- Part 1: Decision-making
- Part 2: Agency
- Part 3: Group membership
- Part 4: Mobility
- Part 5: Relationship dynamics, including IPV
- Part 6: Decision-making on health and nutrition

Module J: Nutrition, Health, and Care of Child
- Part 1: Pregnancy and participation in PSNP4
- Part 2: Use of antenatal and postnatal services
- Part 3: Infant and young child feeding (IYCF) practices
- Part 4: Child health history
- Part 5: Maternal IYCF knowledge and perceptions
- Part 6: Child care activities
- Part 7: Exposure to health and nutrition services
- Part 8: Anthropometry

Module L: Aspirations & Wellbeing
- Part 1: Aspirations

Module K: Participation in VESA groups and SPIR activities

Module L: Aspirations & Wellbeing
- Part 2: Poverty perceptions and wellbeing
- Part 3: Experience with depression and emotional wellbeing

Universal Codes (Include with all CAPI options):

97=Refuse to respond	98=Don't know	99= Not applicable

4.2 Ethics Approval

IFPRI received approval from their Institutional Review Board (IRB) for the SPIR DFSA quantitative evaluation design described in Section 3.

Informed oral consent was collected from all participants prior to the start of the interviews. The entire field team was trained on ethical data collection prior to the start of the field work. Before beginning a household survey, enumerators read the respondent a brief description of the study that was being conducted, informed them that their participation in the study was voluntary and that they could discontinue participating at any time, and asked whether they agreed to respond to the household interview questions. The full written consent can be seen at the beginning of the household survey in Annex I. The enumerator only completed a survey for the household if they received verbal consent to participate in the study from the potential respondent.

Confidentiality of the data is protected by recording survey interview responses using Computer Assisted Personal Interviewing (CAPI), so no hard copy versions of survey questionnaires are available. All files containing raw and analysed data are securely stored in password-protected databases. Access to the complete data is restricted to the IFPRI evaluation team. A unique household ID is assigned to each household. The name and geographic location of the respondent will be kept in a separate data file to which only the research team will have access. Anonymized versions of the data sets that exclude these personal identifiers will be the ones made available for public access.

4.3 Enumeration Teams and Trainings

BDS-CDR, working closely with IFPRI, organized enumeration training for the baseline survey in late January 2018. The two-week training consisted of validating eligibility of households, administering the survey on tablets, and an in-depth training on all modules. The training included one day of pre-testing the survey, on a community of farmers in a rural PSNP community outside of Addis Ababa. A special four-day anthropometry training was held for anthropometry specialists, which included repeated practice measurements of women and children.

Each survey team consisted of three enumerators, one supervisor, and one anthropometric specialist. Each enumerator was expected to complete, on average, three household interviews per day.

4.4 Fieldwork experience

4.4.1 Achieved Sample Size and Interview Completion Rate

The baseline survey enumeration team interviewed a total of 3,396 households out of the target of 3,474 households for the entire sample. Of these, 3,314 household interviews (95.4% of the planned sample) met the inclusion criteria for being in the sample: having a child age 0-35

months and that the child's primary female caregiver (if there is one) is a household member. The 82 additional households appeared to meet the inclusion criteria at the start of the interview, but sometime during the interview it was determined that the household was not eligible to be in the sample. These households were excluded from the analysis, so the final analysis sample includes 3,314 households.

Most of the shortfall from the planned sample of 3,474 households is due to challenges that the survey team had in completing the interview process according to the fieldwork schedule. The survey team was not always able to locate sampled households, confirm their eligibility for the sample (especially the presence of a child age 0-35 months) and arrange for respondents to be available for interviews before the survey team had to leave the village. Nonetheless, the completed sample of 3,314 households provides 17.2 households per kebele on average, which remains within the desired target for statistical power, as described in the Impact Evaluation Inception Report.

Of the 3,314 households in which a primary female respondent was interviewed, 1,920 were from Amhara and 1,394 were from Oromia. All of these households include an index child aged 0-35 months. Interviews with a primary male respondent were completed in 2,756 of the 3,314 households eligible for the sample. Of the 558 households without a primary male respondent interview, 522 (93.5%) were female headed households and most of those would not have head a responsible male (such as a spouse to the female head) eligible to serve as primary male respondent. In only 35 households was a primary male respondent identified but not available for interview.

Table 4.3.1: Summary of household surveys

	Amhara	Oromia	Total
Number of EAs completed	113	80	193
Completed household interviews	1,961	1,435	3,396
Intended household interviews	2,034	1,440	3,474
Number of primary female respondents	1,920	1,394	3,314
Number of primary male respondents	1,464	1,292	2,756
Number of children 0-35 months	1,920	1,394	3,314
Number of children 0-23 months	1,207	874	2,081
T1: L*+ N*	477	330	807
T2: L*+ N	492	365	857
T3: L + N*	501	352	853
C: PSNP only	450	347	797

4.5 Data Quality and Cleaning

Data from the baseline household survey were recorded during the interviews on tablets using a CSPro programme. All data were synced by enumerators (unless there were internet connectivity issues) to a remote server in Dropbox.

BDS and IFPRI were careful to ensure the quality of the data collection. Team supervisors travelled with the enumeration teams, sat in on interviews, and reviewed the data being collected. Enumerators and supervisors conducted revisits to address any issues that may have come up from the data collection teams. Any issues were discussed, and solutions provided. Main problems included duplicate HHIDs, missing data in certain modules, and matching respondents across the household, male, female, and anthropometry surveys.

5. Summarizing Baseline Data: Understanding the Context and Potential for Impact

In this section, we summarize the outcome variables and characteristics of the sample at baseline in order to provide an understanding of the context for the SPIR program and to examine the potential for impact on primary and secondary outcomes. We summarize the data for the entire sample as well as by region, to identify important regional differences that may shape the impact of SPIR.

5.1 Household demographics, child education and housing characteristics

5.1.1 Household demographics

As indicated in Table 5.1.1, the size of the average household in Oromia (6.6 members) is substantially larger than in Amhara (5.1 members). The share of households headed by a female is nearly three times higher in Amhara (27.4%) than in Oromia (9.2%). This rate of female headship in Amhara is consistent with the national average in the 2016 DHS survey (25%) and with the PSNP4 baseline survey for Amhara (32.9%), but the rate of female headship is in the Oromia sample is much lower than in the DHS or in the PSNP4 baseline survey for Oromia (27.4%) (CSA and ICF 2016; Berhane et al. 2016). In the PSNP4 midline survey, the share of PSNP4 beneficiary households that are female headed is only somewhat larger in Amhara (31.6%) than in Oromia (27.5%) (Berhane et al. 2018). It is not clear why female headship is so much lower in Oromia in this sample, but it appears to be consistent with other household demographic characteristics. For example, this is consistent with the share of household heads and primary female respondents that are married (in monogamous marriages) being 15-18 percentage points lower in Amhara than in Oromia.

Many female household heads report that they are married; 30 percent live in a household with an adult male, defined as being older than 17 years. While many of these adults are sons or sons-

in-law, 12 percent of those women who claimed to be the household head also report their spouse in the household roster. Only 15 percent of the primary female respondents indicate that they are mainly engaged in crop production, with a larger share of women in Oromia indicating this activity. Among males, crop production is dominant in employment, with 3 out of 4 male respondents listing crop production as their main activity.

Education attainment by adults is low in the sample, with only 28.6 percent of household heads having any formal education. Primary male respondents in Oromia have slightly more chance of having received formal education than the males in Amhara; in contrast, the primary females in households in Oromia are slightly less likely to have attended formal schooling. As the average primary male or female is in their thirties, they are not in the age cohort that had the opportunity to benefit from the widespread expansion of education opportunities in this century. Currently, 89 percent of children between 6 and 18 are listed as enrolled in school. This pattern of low education attainment of adult household members highlights another dimension of their relative poverty. In the 2016 DHS, the share of women with no education was 54 percent in Amhara and 51 percent in Oromia, while it is 79 percent in Amhara and 81 percent in Oromia in the SPIR baseline sample (CSA and ICF 2016).

5.1.2 Child education

School enrollment is relatively low in the baseline sample, with 68.0 percent of children age 7-18 and 67.4 percent of children age 7-13 currently enrolled in school (Table 5.1.2). Enrollment is 7-8 percentage points higher in Amhara than Oromia. These enrollment rates are lower than reported by UNESCO for Ethiopia on average, which reports a net enrollment rate of 85.4 percent for children of primary school age (7-12) nationally.[12] However, the 2016 DHS survey reports that 71% of children age 7-14 attended school (primary or secondary) at least one day in 2016 (CSA and ICF 2016). This "ever attended" measure for the current school year may be lower than the share of children on the enrollment registers, but it is a reasonable proxy for effective net enrollment (the share of children in an age range enrolled in school and ever attending). That this is close to the enrollment rate in the SPIR baseline data suggests that the SPIR figure may not be lower than expected. Also, the DHS data show virtually no difference in net attendance rates for girls and boys, so we do not disaggregate enrollment by gender.

The average age of starting school is 7.2 on average for children age 7-13 in the sample, which is close to the recommending starting age of 7, suggesting that relatively few enter primary school at a delayed age. Reported attendance rates are high for a measure that captures only long-term absences, with 97.4 percent of children age 7-13 reported to attend at least half time among children who are enrolled in school. Also, the number of days a child age 7-13 was reported to attend school in the past seven days was 4.7 on average. This suggests a school attendance rate of 93.6 percent over the last week if all schools are open five days per week. When school holidays are factored in, reported school attendance in the last week is even higher.

[12] http://uis.unesco.org/country/ET.

5.1.3 Housing characteristics and water sources

Table 5.1.3 summarizes data on housing characteristics and water sources. Virtually all households use solid cooking fuels (e.g., coal, fuelwood or dung). The survey captures detailed information on household water sources. Overall, 62.6-65.2 percent of households report having an improved water source.[13] As a result, 36.4 percent of households report the time needed to fetch water at less than 30 minutes. However, 33.0 percent of households need more than one hour to fetch water and 12.4 percent need more than two hours. Time need to fetch water in Oromia is much higher than in Amhara, with 23.5 percent of Oromia households requiring more than two hours. Also, 76.7 percent of households use the same source for drinking water and other purposes. Virtually no households have an improved toilet. Improved flooring is also rare (6.6%), but 41.9 percent of households have improved roof materials. On average, 56.3 percent of households report having electricity at their dwelling, a figure which is far higher in Oromia (78.9%) than in Amhara (38.1%). The figure for Oromia is surprisingly high, but nearly all of this is driven by reported access to solar power in Oromia. It is not clear whether the solar technology reported is sufficient to provide regular power for electrifying homes or for agricultural or small business uses.

[13] Improved water sources include a household connection (piped), public standpipe, tubewell or borehole, protected well or spring, rainwater collection, or bottled water.

Table 5.1.1: Household demographics

	N	All	Amhara	Oromia
Household size	3,314	5.774	5.154	6.628
		(1.957)	(1.703)	(1.964)
Number of children under the age of 5	3,314	1.419	1.264	1.633
		(0.566)	(0.467)	(0.619)
Female-headed household	3,313	0.196	0.270	0.094
		(0.397)	(0.444)	(0.292)
Age of household head	3,312	38.666	39.353	37.718
		(10.441)	(11.533)	(8.627)
Household head: Married, monogamous	3,310	0.830	0.755	0.933
		(0.376)	(0.430)	(0.250)
Household head: Not married, divorced, widowed, separated	3,310	0.164	0.243	0.056
		(0.371)	(0.429)	(0.230)
Household head has some education	3,311	0.286	0.255	0.328
		(0.452)	(0.436)	(0.470)
Household head has no formal education	3,314	0.714	0.745	0.671
		(0.452)	(0.436)	(0.470)
Household head's main activity is crop production	3,311	0.688	0.664	0.721
		(0.463)	(0.473)	(0.449)
Age of primary female	3,272	30.527	30.426	30.663
		(7.590)	(8.045)	(6.926)
Primary female: Married, monogamous	3,247	0.847	0.780	0.938
		(0.360)	(0.414)	(0.241)
Primary female: Not married, divorced, widowed, separated	3,247	0.151	0.220	0.058
		(0.358)	(0.414)	(0.233)
Primary female has some education	3,248	0.202	0.217	0.181
		(0.401)	(0.412)	(0.386)
Primary female has no formal education	3,248	0.798	0.783	0.819
		(0.401)	(0.412)	(0.386)
Primary female's main activity is crop production	3,248	0.154	0.124	0.195
		(0.361)	(0.330)	(0.396)
Age of primary male	2,750	38.140	38.729	37.475
		(8.887)	(9.137)	(8.549)
Primary male: Married, monogamous	2,744	0.992	0.996	0.988

		(0.087)	(0.064)	(0.107)
Primary male: Not married, divorced, widowed, separated	2,744	0.001	0.001	0.002
		(0.038)	(0.037)	(0.039)
Primary male has some education	2,746	0.336	0.317	0.357
		(0.472)	(0.466)	(0.479)
Primary male has no formal education	2,746	0.664	0.683	0.643
		(0.472)	(0.466)	(0.479)
Primary male's main activity is crop production	2,745	0.773	0.807	0.735
		(0.419)	(0.395)	(0.441)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.1.2: Child education

	N	All	Amhara	Oromia
Children age 7-18 years				
Children 7-18 years currently enrolled in school	6,070	0.680	0.724	0.641
		(0.466)	(0.447)	(0.480)
Age at which children started school	4,492	7.659	7.305	8.011
		(2.153)	(2.084)	(2.164)
Children who attended school at least half the time in the current school year	4,130	0.976	0.969	0.983
		(0.154)	(0.175)	(0.128)
Number of days children attended school in the past seven days	4,022	4.694	4.798	4.584
		(1.074)	(1.081)	(1.056)
Children age 7-13 years				
Children 7-13 years currently enrolled in school	4,582	0.674	0.712	0.641
		(0.469)	(0.453)	(0.480)
Age at which children started school	3,226	7.204	6.918	7.478
		(1.821)	(1.756)	(1.840)
Children who attended school at least half the time in the current school year	3,089	0.974	0.964	0.984
		(0.159)	(0.187)	(0.125)
Number of days children attended school in the past seven days	3,001	4.695	4.801	4.590
		(1.061)	(1.066)	(1.045)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.1.3: Housing characteristics and water sources

	N	All	Amhara	Oromia
Household uses solid cooking fuels	3,312	0.995	0.993	0.998
		(0.071)	(0.085)	(0.046)
Household has improved source of water - rainy season	3,311	0.626	0.736	0.475
		(0.484)	(0.441)	(0.500)
Household has improved source of water - dry season	3,308	0.652	0.748	0.520
		(0.476)	(0.434)	(0.500)
Time taken to fetch water				
Less than 30 mins	3,314	0.364	0.423	0.283
		(0.481)	(0.494)	(0.451)
Between 30 mins-1hr	3,314	0.306	0.358	0.234
		(0.461)	(0.479)	(0.423)
Between 1hr - 2hrs	3,314	0.206	0.176	0.247
		(0.405)	(0.381)	(0.432)
Greater than 2 hours	3,314	0.124	0.043	0.235
		(0.330)	(0.203)	(0.424)
Primary female respondent fetches the water from the source	3,314	0.833	0.869	0.784
		(0.373)	(0.337)	(0.412)
Household uses the same source of drinking water for other purposes	3,062	0.767	0.796	0.731
		(0.423)	(0.403)	(0.443)
Household has improved toilet	3,305	0.007	0.008	0.005
		(0.081)	(0.088)	(0.071)
Toilet facility was built as part of the PSNP Public Works	2,136	0.061	0.031	0.100
		(0.240)	(0.174)	(0.301)
Household has improved roof material	3,312	0.419	0.405	0.440
		(0.494)	(0.491)	(0.497)
Household has improved floor material	3,307	0.066	0.110	0.005
		(0.248)	(0.313)	(0.071)
Number of bedrooms	3,312	1.528	1.495	1.574
		(2.283)	(2.450)	(2.030)
Household has electricity, mains	3,064	0.116	0.202	0.009
		(0.320)	(0.401)	(0.097)
Household has electricity, generator	3,064	0.003	0.004	0.001
		(0.051)	(0.059)	(0.038)

Household has electricity, solar panel	3,064	0.390	0.110	0.735
		(0.488)	(0.313)	(0.441)
Household has electricity, other	3,064	0.055	0.066	0.042
		(0.228)	(0.248)	(0.201)
Household has no electricity	3,064	0.437	0.619	0.211
		(0.496)	(0.486)	(0.408)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

5.2 Participation in the PSNP and start of VESA group formation

We summarize households' exposure to the PSNP and participation in PSNP activities over time in Table 5.2.1. The table confirms that all households in the sample for this impact evaluation study have at least one member currently participating in PSNP4. Past participation in PSNP Public Works is high over the last two years. In both regions, roughly 84 percent of PSNP4 households had been in the PW part of the program two years earlier (Tir 2008 – Tahisas 2009 in the Ethiopian Calendar, EC), during the first year of the PSNP4, under entirely new targeting. The results show that the process of re-targeting for PSNP4 worked very differently in the two regions. Two years before the start of PSNP4, in the period Tir 2006 – Tahisas 2007 (January-December 2014), 54.1 percent of current PSNP4 households in Amhara had been in PW, while only 14.1 percent of current PSNP4 households in Oromia were in PW at that time. The pattern of persistence in eligibility for the PSNP in Amhara between Phase 3 and Phase 4 of the PSNP is consistent with prior experience, while in Oromia it appears that retargeting led to many new households being selected to participate in the PSNP4.

Coverage of Direct Support varied by region, with 19.8 percent of PSNP4 beneficiaries in Amhara receiving DS payments in the past year and 3.9 percent of PSNP4 beneficiaries in Oromia receiving DS payments in the past year. This may include both Permanent Direct Support (PDS) provided to labor-scarce households, including the elderly and disabled, and Temporary Direct Support (TDS) provided to women who are pregnant or within 12 months post partum. However, women who receive TDS must be living in PW client households before being transitioned to receive TDS, so these households may identify themselves as participants in PW, TDS or both. The low share of DS recipients in Oromia is consistent with the low share of female-headed household in Oromia, which may be more likely to qualify for DS than male-headed households. Overall, we would expect a low share of PDS households because they are less likely to have a child under the age of 3, one of the eligibility criterion to be in the study sample. The share of current DS households that received DS payments before the start of PSNP4 is also much higher in Amhara than in Oromia.

Other evidence on household behaviour around the PSNP includes that household heads alone made decisions about who in the household would work on PSNP for 63.2 percent of households. In addition, few households sold the food they received from PW (3.4%) or from DS (1.0%). Most households were aware of the Livelihoods Component of PSNP4 (73%) and 40.2 percent of households had joined a Livelihoods Group.

Exposure to SPIR intervention activities at the time of the baseline survey is summarized in Table 5.2.2. A small share of households (5.9%) had already joined a VESA group, one of the first SPIR activities in project communities. This small level of program initiation pre-baseline is not likely to have affected outcomes before the baseline survey other than possibly some measures of financial literacy and some nutrition knowledge indicators. In section 6, we test whether means of these indicators are balanced across treatment and control groups, so we can observe whether this limited, early initiation had any effect on the study design.

An objective of the PSNP4 is to achieve meaningful graduation from the program for some participants, where their incomes would have improved sufficiently that they are no longer at substantial risk of future food insecurity or poverty. Survey respondents were asked about their awareness of graduation criteria and their own experience with graduation or other terminations from the PSNP4. These responses are summarized in Table 5.2.3. Most respondents list household income as an important criterion determining graduation from the program. When livestock holdings and other assets are included, measures of income and wealth were mentioned by 88.0 percent of respondents. Other criteria mentioned include months of food insecurity or that graduation decisions are arbitrary. It is interesting that 6.9 percent of respondents describe themselves as having graduated from PSNP4 in the last two years. These households indicated they are currently PSNP4 clients, so they must have left the program and rejoined, or else they are receiving their final payments before being removed from the programme. Of those that had graduated at some point in the last two years, 58.8 percent indicated that this was based on their income level; 16.9 percent said it was because of their level of food insecurity. Seventy-two households indicated that they had "self-graduated" from PSNP4 in the last two years, but nearly all of these indicated they had graduated too early, which would explain their current participation in PSNP4. Nonetheless, the households that reported having graduated from the PSNP4 in the last two years were visibly better off, with significantly higher asset holdings than other PSNP4 clients in the sample.

Table 5.2.1: Access to the PSNP

	N	All	Amhara	Oromia
Household has at least one member currently participating in PSNP	3,314	1.000	1.000	1.000
		(0.000)	(0.000)	(0.000)
Household participated in PNSP PW activities between TIR 2006 and TAHISAS 2007	3,313	0.374	0.541	0.145
		(0.484)	(0.498)	(0.352)
Household participated in PNSP PW activities between TIR 2007 and TAHISAS 2008	3,313	0.582	0.731	0.377
		(0.493)	(0.443)	(0.485)
Household participated in PNSP PW activities between TIR 2008 and TAHISAS 2009	3,313	0.845	0.851	0.836
		(0.362)	(0.356)	(0.370)
Household participated in PNSP PW activities between TIR 2009 and TAHISAS 2010	3,313	0.914	0.867	0.978
		(0.280)	(0.340)	(0.145)
Household head solely made the decision about who would work on PSNP public works	3,174	0.632	0.679	0.570
		(0.482)	(0.467)	(0.495)
Household head solely made the decision about how the PW transfers were to be used	3,313	0.493	0.549	0.415
		(0.500)	(0.498)	(0.493)
Households that sold some food received as PW payments for cash	2,406	0.034	0.034	0.032
		(0.180)	(0.182)	(0.175)
Household received Direct Support payments between TIR 2006 and TAHISAS 2007	3,313	0.053	0.087	0.006
		(0.224)	(0.281)	(0.080)
Household received Direct Support payments between TIR 2007 and TAHISAS 2008	3,313	0.087	0.142	0.011
		(0.282)	(0.349)	(0.107)
Household received Direct Support payments between TIR 2008 and TAHISAS 2009	3,313	0.121	0.188	0.029
		(0.327)	(0.391)	(0.169)
Household received Direct Support payments between TIR 2009 and TAHISAS 2010	3,313	0.131	0.198	0.039
		(0.338)	(0.399)	(0.195)
Households that sold some food received as Direct Support payments for cash	1,572	0.010	0.009	0.011
		(0.097)	(0.093)	(0.106)
Household head solely made the decision about how the Direct Support transfers w	2,015	0.400	0.500	0.290
		(0.490)	(0.500)	(0.454)
Household received government transfers NOT related to PSNP between TIR 2009 and	3,313	0.149	0.095	0.222
		(0.356)	(0.294)	(0.416)
Household has heard about new PSNP 'Livelihoods Component'	3,313	0.730	0.776	0.665
		(0.444)	(0.417)	(0.472)
Household has joined a new PSNP Livelihoods Group	2,417	0.402	0.420	0.373
		(0.490)	(0.494)	(0.484)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.2.2: Exposure to SPIR activities

	N	All	Amhara	Oromia
Household is a member of a VESA group	2,828	0.059 (0.236)	0.088 (0.283)	0.027 (0.161)
Household part of a VESA group and with a child under 2 years of age received counseling	168	0.417 (0.494)	0.451 (0.499)	0.286 (0.458)
Household part of a VESA group, participated in 2 weeks of food demonstration session	168	0.339 (0.475)	0.376 (0.486)	0.200 (0.406)
Household participated in VESA group discussions about child, maternal and adolescent nutrition and WASH behaviors	168	0.500 (0.501)	0.571 (0.497)	0.229 (0.426)
Household part of a VESA group participated in Public World group counseling session	168	0.649 (0.479)	0.699 (0.460)	0.457 (0.505)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.2.2a: Access to the PSNP and SPIR activities, by woreda, Amhara

	Bugna	Dahena	Gaz Gibla	Lasta	Meket	Sekota	Wadla
Household is a member of a VESA group	0.000 (0.000)	0.010 (0.102)	0.012 (0.108)	0.009 (0.097)	0.176 (0.382)	0.072 (0.259)	0.083 (0.277)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.2.2b: Access to the PSNP and SPIR activities, by woreda, Oromia

	Chiro	Gemechis	Grawa	Siraro	Daro Lebu	Kurfachelle
Household is a member of a VESA group	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.031 (0.173)	0.000 (0.000)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.2.3: Graduation from the PSNP

	N	All	Amhara	Oromia
Household perceptions of the criteria for graduation from the PSNP				
based on income	3,027	0.689	0.698	0.679
		(0.463)	(0.459)	(0.467)
based on livestock	3,027	0.101	0.127	0.070
		(0.301)	(0.333)	(0.256)
based on other assets	3,027	0.090	0.107	0.069
		(0.286)	(0.309)	(0.253)
months of food insecurity	3,027	0.037	0.006	0.074
		(0.189)	(0.074)	(0.262)
based on skills	3,027	0.010	0.006	0.014
		(0.097)	(0.078)	(0.116)
based on advice of community leader	3,027	0.009	0.004	0.015
		(0.096)	(0.065)	(0.122)
self-graduation	3,027	0.006	0.006	0.006
		(0.079)	(0.078)	(0.080)
none/arbitrary	3,027	0.017	0.010	0.024
		(0.127)	(0.102)	(0.152)
based on other criteria	3,027	0.042	0.035	0.049
		(0.200)	(0.184)	(0.217)
Household described itself as having graduated from the PSNP in the last 2 years	3,184	0.069	0.034	0.115
		(0.254)	(0.181)	(0.319)
Among households that graduated in the last 2 years, reasons for graduating				
based on income	221	0.588	0.918	0.463
		(0.493)	(0.277)	(0.500)
based on livestock	221	0.077	0.049	0.087
		(0.267)	(0.218)	(0.283)
based on other assets	221	0.036	0.000	0.050
		(0.187)	(0.000)	(0.219)
months of food insecurity	221	0.163	0.000	0.225
		(0.370)	(0.000)	(0.419)
skills	221	0.054	0.000	0.075
		(0.227)	(0.000)	(0.264)
advice of community leader	221	0.009	0.000	0.013
		(0.095)	(0.000)	(0.111)
self-graduation	221	0.018	0.000	0.025

		(0.134)	(0.000)	(0.157)
no reason/arbitrary	221	0.023	0.016	0.025
		(0.149)	(0.128)	(0.157)
other	221	0.032	0.016	0.037
		(0.176)	(0.128)	(0.191)
Household described itself as having self-graduated from the PSNP	221	0.326	0.279	0.344
		(0.470)	(0.452)	(0.476)
Household thought they graduated too early	215	0.312	0.382	0.287
		(0.464)	(0.490)	(0.454)
Household stopped participating in PSNP without graduating in last 2 years	3,105	0.030	0.029	0.030
		(0.170)	(0.168)	(0.171)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

5.3 Child nutrition and feeding practices

5.3.1 Child nutritional status

We describe several key indicators of nutritional status among index children 0-35 months old: (1) height- (or length-) for-age z-score (HAZ), (2) weight-for-height (or length) z-score (WHZ), and (3) mid-upper arm circumference (MUAC). Height-for-age and weight-for-height are normed against the WHO 2006 reference population to compare against a benchmark of children the same age and sex (WHO 2006); a "Z-score" denotes how many standard deviations above or below the reference population's median a particular child's measurements are.

Each measure captures a distinct but related dimension of nutritional status. Height-for-age (also referred to as "length-for-age" for children under age 2, since they are measured lying down) captures a child's cumulative nutritional environment in terms of diet and infection. Therefore, low height-for-age is a measure of chronic undernutrition. Evidence indicates that interventions must occur intensively and with long duration during the first 1,000 days in order to affect height-for-age, with very little potential for remediation later (Victora et al. 2010). Height-for-age is a strong predictor of adult health, education, and labor outcomes and is understood to capture early development (Hoddinott et al. 2013; Black et al. 2013). A child with height-for-age less than 2 standard deviations below the mean of the WHO 2006 reference population (HAZ <-2) is referred to as "stunted." Weight-for-height captures short-term nutritional status; low weight-for-height is considered a measure of acute malnutrition. Unlike height, weight can be changed relatively quickly through changes in the nutrition environment and can be affected beyond the first 1,000 days. Because a child's weight depends on height, the weight-for-height measure allows distinguishing low weight-for-age driven by low height-for-age from low weight-for-age, given height-for-age. A child with weight-for-height less than 2 standard deviations below the mean for the WHO 2006 reference population (WHZ <-2) is referred to as "wasted." MUAC is used to detect moderate acute malnutrition (MAM) and severe acute malnutrition (SAM). Children with SAM are at an elevated risk of death and should receive a therapeutic diet. The cutoff for MAM is between 11.5-12.5 cm and for SAM is less than 11.5 cm.

Table 5.3.1 provides descriptive statistics for the index children age 0-35 months in the study and disaggregates by region. The average height-for-age of children in the sample is 1.401 standard deviations below the median for the reference population; the average weight-for-height is 0.505 standard deviations below the median for the reference population; and the average MUAC is 13.34 cm. These averages translate to 37.4 percent of children being stunted, 13.8 percent being wasted, and 6.2 percent having severe acute malnutrition. These are high rates of chronic and acute undernutrition. The prevalence of wasting is particularly concerning, given that it is almost 4 percent higher in this sample than in the 2016 DHS (at 10%). According to the WHO Crisis classification table, 14 percent of child wasting denominates the nutritional situation as *serious*, whereas from 15 percent onward the situation becomes *critical* (WHO 2000). Across regions, Amhara has lower HAZ, WHZ, and MUAC leading to a larger proportion of children being stunted (42.7 percent versus 30.3 percent), wasted (15.5 percent versus 11.5 percent), and having

severe acute malnutrition (7.7 percent versus 4.2 percent). Rates of stunting are slightly lower in both Amhara and Oromia in the study sample compared to the 2016 Ethiopian Demographic Health Survey (EDHS), where stunting is reported to be 46 percent in Amhara and 37 percent in Oromia (CSA and ICF 2016). This likely reflects that the SPIR baseline survey was two full years after the 2016 DHS, during a period when stunting prevalence has been steadily declining.

The age profile of HAZ and WHZ for children age 0-35 months is shown in Figures 5.3.1 and 5.3.2. The age profile of HAZ reveals a steep decline in mean z-score until around 18-20 months and then increasing slightly and stabling out. Children in Amhara have lower HAZ than children in Oromia across all ages. The age profile of WHZ reveals a steep decline in mean z-score until around 12-15 months, then an increase in the average z-score in Oromia, while Amhara continues to decrease until around 28 months. These growth faltering patterns indicate the need to focus nutrition interventions in the first two years of life (Victora et al. 2010).

Figure 5.3.1 Age profile of height-for-age z-scores, children 0-35 months

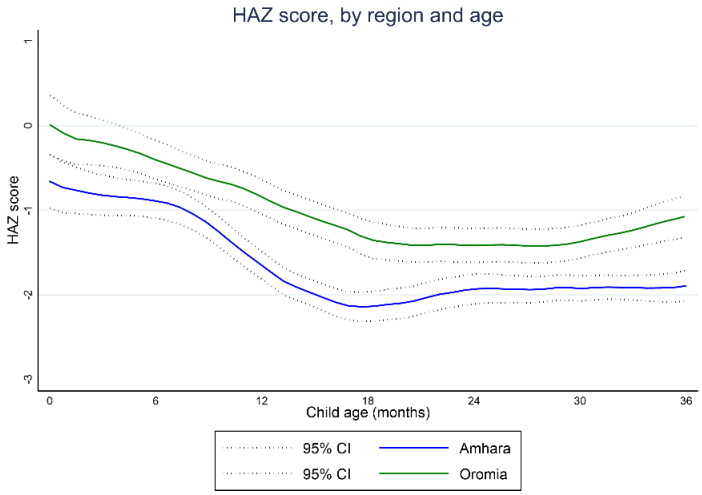
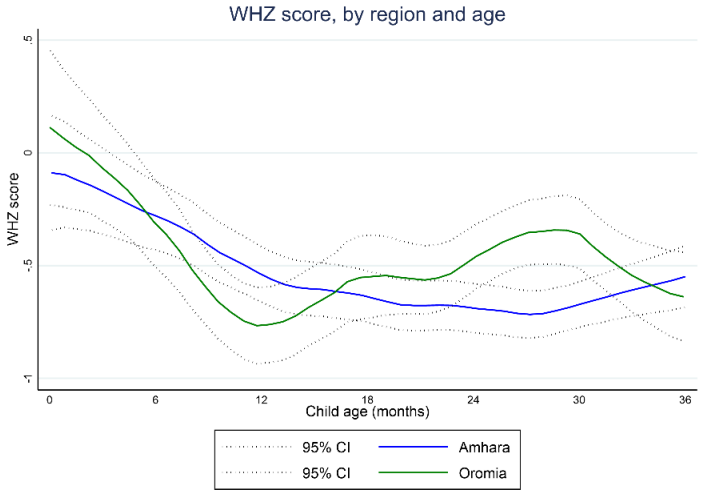


Figure 5.3.2 Age profile of weight-for-height z-scores, children 0-35 months



5.3.2 Infant and young child feeding practices

Infant and young child feeding (IYCF) practices are important determinants of children's nutritional status under two years of age. IYCF indicators are composed of seven core indicators: (1) Early initiation of breastfeeding defined as the "Proportion of children born in the last 24 months who were put to the breast within one hour of birth"; (2) Exclusive breastfeeding under 6 months defined as the "Proportion of infants 0–5 months of age who are fed exclusively with breast milk"; (3) Continued breastfeeding at 1 year defined as the "Proportion of children 12–15 months of age who are fed breast milk"; (4) Introduction of solid, semi-solid or soft foods defined as the "Proportion of infants 6–8 months of age who receive solid, semi-solid or soft foods"; (5) Minimum dietary diversity defined as the "Proportion of children 6–23 months of age who receive foods from 4 or more food groups"; (6) Minimum meal frequency defined as the proportion of children 6–23 months of age who receive solid, semi-solid, or soft foods the minimum number of times; and (7) Minimum acceptable diet which combines the minimum dietary diversity and minimum meal food frequency (WHO 2008). With the exception of the first indicator, each indicator is relevant only to children in a particular age interval and rely on mother's report of current rather than recalled practices.

Table 5.3.2 reveals high rates of early initiation of breast-feeding (84.5 percent) and almost universal continued breastfeeding (99.5 percent). Exclusive breastfeeding under 6 months is slightly lower at 73.1percent. Among the subset of children age 6-8 months at baseline, only about half (45.2 percent) had been introduced to solid, semi-solid, or soft foods, implying that a delay in appropriate complementary feeding is common. Although 43.8 percent of children age 6-23 months achieved minimum meal frequency, only 1.8 percent achieved minimum dietary diversity (consuming at least 4 of the 7 food groups defined by the WHO), compared with 14 percent in the Ethiopia DHS. As a result, the share of children age 6-23 months that received a

minimum acceptable diet is also very low at 1 percent. Across regions there is relatively little variation in means, with the exception of the proportion of children 6-8 months that received solid, semi-solid, or soft food which was nearly double in Oromia compared to Amhara (63.4 percent versus 33.8 percent).

To better understand the food consumption patterns of children and the low rates achieving minimum dietary diversity, Table 5.3.3 presents the proportion of children consuming each of the 7 food groups in the last 24 hours, in addition to an indicator for whether the child consumed any animal sourced food. Most children consumed grains, roots, and tuber, and approximately a quarter of children consumed legumes and nuts. However, consumption of dairy, flesh foods, eggs, and fruits and vegetables are very low with 10 percent or lower consuming each food group. Regional differences also emerge, with a larger share of children in Amhara consuming legumes and nuts (33 percent in Amhara versus 14 percent in Oromia), and a lower share of children consuming dairy (4 percent in Amhara versus 19 percent in Oromia). Overall the proportion of children eating any animal sourced foods is 14.5 percent, with large differences across Amhara and Oromia (8 percent in Amhara and 24 percent in Oromia).

5.3.3 Mother and father nutrition knowledge

It is hypothesized that caregivers (both women and men) who have been exposed to the household level counselling on maternal nutrition and IYCF, will have increased their knowledge on these topics. These practices include appropriate breastfeeding, the timely introduction of (semi-) solid foods at the age of 6 months, the number of feedings, feeding a sick child, the inclusion of nutrient dense foods into the complementary diet and appropriate WASH practices. Given that men will also be targeted in the N* treatment group, we assess both the primary male and primary female's baseline knowledge on nutrition and IYCF.

In total 25 questions with respect to nutrition, IYCF and WASH were asked to women and the same 22 out of 25 questions were asked to men. The three questions dropped for men were specific to a mother's colostrum, what she should do if she thinks the baby is not getting enough breastmilk, and special foods a mother can make to complement breastmilk. Eleven multiple response option questions for women and 9 for men are also dropped from this analysis due to programming errors in the CAPI, thus reducing the number of questions used to create total scores to 14 and 13.

Table 5.3.4 reveals that on average mothers answered 7.2 out of 14 questions correctly for a score of 51.41 percent, while fathers answered 6.27 out of 13 questions correctly for a score of 48.21. Mothers and fathers have higher scores in Oromia compared to Amhara. For mothers, questions with the highest scores were on breastfeeding while questions with the lowest scores were on complementary feeding and, in particular, the number of times a child at different ages should eat (Table 5.3.5). A similar pattern emerges for fathers' knowledge (5.3.6).

Table 5.3.1: Child Anthropometry (age 0-35 months)

	N	All	Amhara	Oromia
Height-for-age z-score(HAZ)	3,195	-1.401	-1.664	-1.045
		(1.871)	(1.810)	(1.894)
Proportion stunted (HAZ<-2SD)	3,195	0.374	0.427	0.303
		(0.484)	(0.495)	(0.460)
Weight-for-height z-score(WHZ)	3,176	-0.505	-0.543	-0.453
		(1.497)	(1.561)	(1.404)
Proportion wasted (WHZ<-2SD)	3,176	0.138	0.155	0.115
		(0.345)	(0.362)	(0.319)
Proportion with moderate acute malnutrition (-3 SD<=WHZ<-2SD)	3,176	0.082	0.093	0.067
		(0.275)	(0.291)	(0.250)
Proportion with severe acute malnutrition (WHZ<-3 SD)	3,176	0.057	0.063	0.049
		(0.232)	(0.242)	(0.216)
Mid-upper arm circumference (MUAC)	3,286	13.336	13.216	13.501
		(1.384)	(1.463)	(1.249)
Proportion with moderate acute malnutrition (11.5 cm<=MUAC<12.5 cm)	3,286	0.161	0.171	0.148
		(0.368)	(0.377)	(0.355)
Proportion with severe acute malnutrition (MUAC<11.5)	3,286	0.062	0.077	0.042
		(0.242)	(0.267)	(0.202)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.3.2: Infant and Young Child Feeding Practices

	N	All	Amhara	Oromia
Children born in the last 24 months who were put to the breast within one hour	1,853	0.845	0.837	0.857
		(0.362)	(0.369)	(0.350)
Infants 0-5 months of age who are fed exclusively with breast milk	413	0.731	0.753	0.699
		(0.444)	(0.432)	(0.460)
Children 12-15 months of age who are fed breast milk	371	0.995	1.000	0.986
		(0.073)	(0.000)	(0.119)
Infants 6-8 months of age who receive solid, semi-solid or soft foods	261	0.452	0.338	0.634
		(0.499)	(0.474)	(0.484)
Children 6-23 months of age who meet the minimum dietary diversity	1,440	0.018	0.011	0.028
		(0.133)	(0.106)	(0.166)
Children 6-23 months of age who meet the minimum meal frequency	1,440	0.438	0.434	0.442
		(0.496)	(0.496)	(0.497)
Children 6-23 months of age who receive a minimum acceptable diet	1,440	0.010	0.008	0.012
		(0.098)	(0.089)	(0.111)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.3.3: Food groups consumed in the last 24 hours by children 6-23 months

	N	All	Amhara	Oromia
Consumed grains, roots and tubers	1,440	0.759	0.759	0.758
		(0.428)	(0.428)	(0.428)
Consumed legumes and nuts	1,440	0.257	0.333	0.139
		(0.437)	(0.472)	(0.346)
Consumed dairy products	1,440	0.101	0.042	0.194
		(0.302)	(0.201)	(0.395)
Consumed flesh foods	1,440	0.014	0.015	0.012
		(0.117)	(0.121)	(0.111)
Consumed eggs	1,440	0.047	0.039	0.060
		(0.212)	(0.193)	(0.238)
Consumed vitamin A fruits and vegetables	1,440	0.047	0.030	0.073
		(0.211)	(0.170)	(0.260)
Consumed other fruits and vegetables	1,440	0.022	0.014	0.036
		(0.147)	(0.116)	(0.185)
Consumed animal sourced food	1,440	0.145	0.083	0.242
		(0.352)	(0.276)	(0.428)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.3.4: Infant and Young Child Feeding (IYCF) Knowledge

	N	All	Amhara	Oromia
Maternal IYCF knowledge score (0-14)	3,314	7.198	6.754	7.809
		(2.258)	(2.347)	(1.973)
Maternal IYCF knowledge score (percent)	3,314	51.412	48.240	55.780
		(16.130)	(16.767)	(14.091)
Male IYCF knowledge score (0-13)	2,756	6.268	5.960	6.617
		(2.069)	(2.051)	(2.034)
Male IYCF knowledge score (percent)	2,756	48.214	45.844	50.899
		(15.913)	(15.777)	(15.644)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Missing responses on any question were treated as an incorrect response.

Table 5.3.5: IYCF Knowledge Questions, primary female

	N	All	Amhara	Oromia
Knows how long after birth baby should start breastfeeding	3,314	0.896	0.889	0.905
		(0.306)	(0.314)	(0.294)
Knows what a mother should do with the colostrum	3,314	0.843	0.851	0.832
		(0.364)	(0.356)	(0.374)
Knows what age a baby should be exclusively breastfed until	3,314	0.864	0.847	0.888
		(0.343)	(0.360)	(0.315)
Knows the age at which a baby should first start to receive liquids other than breastmilk	3,314	0.490	0.433	0.568
		(0.500)	(0.496)	(0.496)
Knows the age at which a baby should first start to receive foods in addition to breastmilk	3,314	0.337	0.278	0.420
		(0.473)	(0.448)	(0.494)
Knows which seasoning is fortified with iodine	3,314	0.761	0.704	0.839
		(0.427)	(0.457)	(0.367)
Knows the common problem with gruels given as first foods to babies	3,314	0.359	0.304	0.435
		(0.480)	(0.460)	(0.496)
Knows that a 1-year old child cannot eat alone without any supervision	3,314	0.866	0.836	0.907
		(0.341)	(0.370)	(0.291)
Knows how many times a day a 6-8 month old baby that is still breastfeeding should eat	3,314	0.106	0.103	0.109
		(0.307)	(0.304)	(0.312)
Knows how many times a day a 9-11 month old baby that is still breastfeeding should eat	3,314	0.240	0.264	0.208
		(0.427)	(0.441)	(0.406)
Knows how many times a day a 12-24 month old child should eat (excluding breast	3,314	0.336	0.309	0.373
		(0.472)	(0.462)	(0.484)
Knows how often a baby 6-23 months old should eat animal source foods	3,314	0.288	0.233	0.364
		(0.453)	(0.423)	(0.481)
Knows how much a child should be fed when sick	3,314	0.396	0.335	0.481
		(0.489)	(0.472)	(0.500)
Knows how often a child should be fed when sick	3,314	0.416	0.368	0.481
		(0.493)	(0.482)	(0.500)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Missing responses on any question were treated as an incorrect response.

Table 5.3.6: IYCF Knowledge Questions, primary male

	N	All	Amhara	Oromia
Knows how long after birth baby should start breastfeeding	2,756	0.781	0.775	0.787
		(0.414)	(0.418)	(0.409)
Knows what age a baby should be exclusively breastfed until	2,756	0.799	0.748	0.857
		(0.401)	(0.434)	(0.350)
Knows the age at which a baby should first start to receive liquids other than breastmilk	2,756	0.484	0.443	0.532
		(0.500)	(0.497)	(0.499)
Knows the age at which a baby should first start to receive foods in addition to breastmilk	2,756	0.331	0.277	0.393
		(0.471)	(0.447)	(0.489)
Knows which seasoning is fortified with iodine	2,756	0.796	0.762	0.835
		(0.403)	(0.426)	(0.371)
Knows the common problem with gruels given as first foods to babies	2,756	0.319	0.251	0.395
		(0.466)	(0.434)	(0.489)
Knows that a 1-year old child cannot eat alone without any supervision	2,756	0.901	0.938	0.859
		(0.299)	(0.242)	(0.348)
Knows how many times a day a 6-8 month old baby that is still breastfeeding should eat	2,756	0.103	0.097	0.111
		(0.305)	(0.296)	(0.314)
Knows how many times a day a 9-11 month old baby that is still breastfeeding should eat	2,756	0.251	0.283	0.214
		(0.434)	(0.451)	(0.410)
Knows how many times a day a 12-24 month old child should eat (excluding breast	2,756	0.348	0.336	0.361
		(0.476)	(0.473)	(0.481)
Knows how often a baby 6-23 months old should eat animal source foods	2,756	0.301	0.272	0.334
		(0.459)	(0.445)	(0.472)
Knows how much a child should be fed when sick	2,756	0.414	0.377	0.456
		(0.493)	(0.485)	(0.498)
Knows how often a child should be fed when sick	2,756	0.439	0.400	0.484
		(0.496)	(0.490)	(0.500)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Missing responses on any question were treated as an incorrect response.

5.4 Household food security, dietary diversity and consumption

5.4.1 Household food security; household and women's dietary diversity

A primary objective of the SPIR project is to improve household food security. We examine two dimensions of food security captured in the baseline survey. The first is the food gap, a count of the number of months in the last 12 months that the household had trouble meeting its food needs. This measure is sometimes interpreted as the length of the lean season. It is a commonly used measure of food security in discussions around targeting and effectiveness of the PSNP. Table 5.4.1 shows that the average food gap in the baseline survey was 2.1 months, which is a modest figure historically for poor households in Ethiopia, reflecting that EC 2009-2010 was not a very difficult year for food security on average, though many households suffered substantial food insecurity. The food gap is much higher in Oromia (3.3 months) than in Amhara (1.4 months). This stands in contrast to the regional differences in nutritional status highlighted above, where stunting and severe wasting are higher in Amhara than in Oromia.

The second of food security is a food security index constructed based on responses to 9 questions asked to the primary female respondent about household behaviors related to coping strategies for food insecurity. Six of the questions ask about coping strategies the household used during the worst month of food insecurity in the last 12 months: (i) number of times per day adults ate, (ii) number of times per day children ate, (iii) number of times per day adolescents ate, (iv) whether household members consumed less preferred foods, (v) whether household members consumed wild foods, and (vi) whether household members consumed seed stocks. Three other questions asked about the number of times adults, children and adolescents ate during a good month. Using responses to these 9 questions, we constructed a food security index using principal components analysis (PCA), a method of constructing an index from a set of potentially correlated variables that adjusts for the degree of similarity in the questions in order to construct a meaningful index. We then constructed three indicators for levels of food insecurity based on three quantiles (tertiles) of the food security index. The regional distribution of these variables is summarized in Table 5.4.1. For the sample overall, roughly one third of the sample resides at each level of this food security index tertiles, by construction. However, the regional differences are informative. In Oromia, 37.1 percent of households are in the most food insecure group, while in Amhara only 21.7 percent of households are among the most food insecure. In Oromia, nearly half of the sample is in the middle food insecurity group, and only 16.7 percent of households are among the most food secure. In Amhara, nearly half of the sample is in the most food secure group.

Table 5.4.1 also reports average household dietary diversity based on a count of the number of food groups out of 12 that household members consumed foods from in the past 7 days. Average household dietary diversity was 4.6 food groups, with only a small difference between Amhara and Oromia. By this household dietary diversity score (HDDS) measure, Oromia has slightly better dietary diversity. Finally, we report a measure of women's dietary diversity from the WHO that counts the number of food groups (out of 10) that capture all of the foods consumed

by the primary female respondent in the last 24 hours. The mean for this women's dietary diversity score (WDDS) is low, at 2.05. By this measure, women in Oromia are substantially worse off (at 1.88) than those in Amhara (at 2.16). The distributions of these two dietary diversity measures suggest intrahousehold inequality in dietary diversity across the two regions, with women somewhat more disadvantaged in Oromia than in Amhara with regard to their dietary diversity compared to the rest of the household.[14]

5.4.2 Household consumption and poverty

The baseline survey gathered data on household food consumption in the past 7 days and nonfood consumption in the past month that we used to develop measures of total, food and nonfood consumption at the household level and per adult equivalent in the past month. We also used the food consumption data to develop estimates of the value of food consumed in the past month. Table 5.4.2 summarizes these data. The table shows that the average value of household consumption in the past month was 2,380 Birr. Total household consumption was slightly higher in Oromia than in Amhara. The value of mean consumption per adult equivalent in the past month was 591 Birr and was lower in Oromia (520 Birr) than in Amhara (643 Birr) due, in part, to larger households in Oromia. This is consistent with evidence from the PNSP4 midline report showing that households in Oromia were poorer on average than those in Amhara (Berhane et al. 2018). Figure 5.4.1 shows the distribution of the natural log of consumption per adult equivalent in the last month in each region. The distribution in Oromia is nearly everywhere to the left of the distribution in Amhara. The vertical red line in Figure 5.4.1 represents the international poverty line for extreme poverty ($1.25 per person per day). We can compare the figure to estimates in Table 5.4.2 of the poverty rate (below $1.90 per person per day) and the extreme poverty rate (below $1.25 per person per day) which are 45.3 percent and 22.2 percent, respectively. At 30.0 percent, the extreme poverty rate is much higher in Oromia than in Amhara, at 16.5 percent, which is consistent with Figure 5.4.1.

The value of food consumption per adult equivalent was 473 Birr and the value of nonfood consumption per adult equivalent was 119. This suggests that food represents 80 percent of the value of monthly consumption; this share may be overstated because the food consumption module is more detailed than what is captured in nonfood consumption. Table 5.4.2 also reports mean calorie consumption per adult equivalent per day at 2,554 calories, but this variable is very noisy as seen by the high standard deviation in the table. Calorie consumption per adult equivalent is slightly higher in Oromia than in Amhara. Figure 5.4.2 shows the distribution of the natural log of calorie consumption per adult equivalent in each region. The distribution in Oromia is generally to the right of the distribution in Amhara.

[14] A caveat to this conclusion about intrahousehold inequality in dietary diversity is that the HDDS and WDDS are constructed using different data sources. The HDDS is constructed from the food consumption module and the WDDS is constructed from a module designed for this purpose.

Figure 5.4.1: Distribution of monthly consumption expenditure per adult equivalent

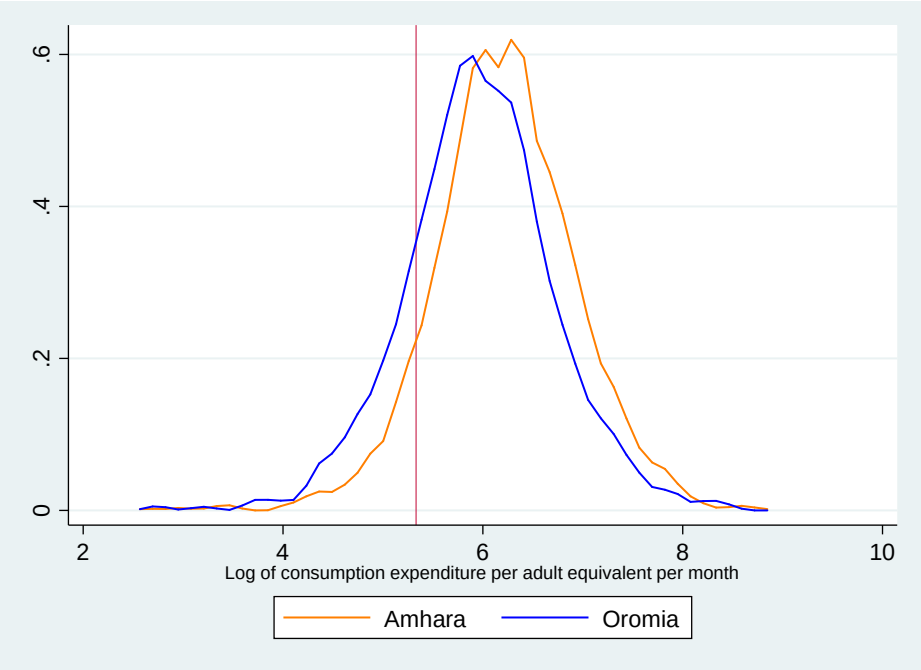


Figure 5.4.2: Distribution of daily calorie consumption per adult equivalent

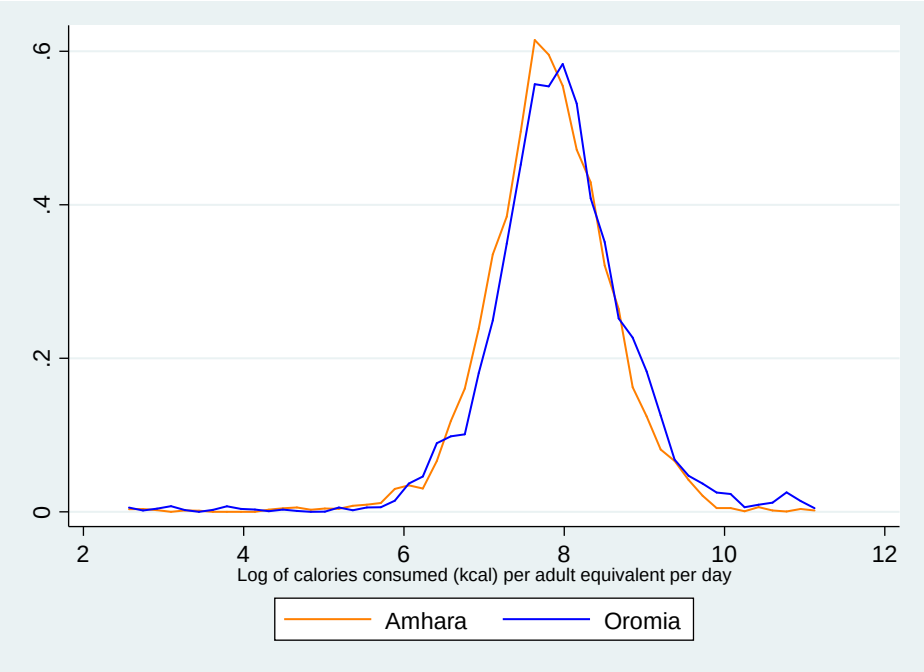


Table 5.4.1: Food security and dietary diversity

	N	All	Amhara	Oromia
Food gap in months (0-12)	3,305	2.187	1.355	3.331
		(2.414)	(1.589)	(2.849)
Food security index: First tertile	3,314	0.281	0.216	0.371
		(0.450)	(0.411)	(0.483)
Food security index: Second tertile	3,314	0.364	0.290	0.466
		(0.481)	(0.454)	(0.499)
Food security index: Third tertile	3,314	0.355	0.494	0.164
		(0.479)	(0.500)	(0.370)
Number of food groups (of 12) the household consumed in the past 7 days	3,314	4.621	4.481	4.813
		(1.377)	(1.249)	(1.516)
Number of food groups (of 10) women consumed the previous day or night	3,314	2.046	2.163	1.884
		(1.185)	(1.108)	(1.267)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.4.2: Consumption expenditure

	N	All	Amhara	Oromia
Consumption expenditure in the past month (Birr)	3,313	2,380	2,345	2,428
		(2,013)	(1,877)	(2,186)
Consumption expenditure per adult equivalent in the past month (Birr)	3,313	591	643	520
		(535)	(531)	(533)
Food consumption expenditure per adult equivalent in the past month (Birr)	3,313	473	520	407
		(482)	(480)	(477)
Nonfood consumption expenditure per adult equivalent in the past month (Birr)	3,313	119	123	112
		(150)	(140)	(163)
Calories (kcal) of food consumption per adult equivalent per day	3,314	2,554	2,438	2,739
		(10,831)	(3,396)	(16,183)
Share of the population living in households with consumption per person below the $1.90 poverty line	3,314	0.453	0.384	0.547
		(0.498)	(0.486)	(0.498)
Share of the population living in households with consumption per person below the $1.25 poverty line	3,314	0.222	0.165	0.300
		(0.416)	(0.371)	(0.458)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Median values for calories consumed are reported.

5.5 Livelihood outcomes: assets, financial inclusion, aspirations, agricultural production and exposure to shocks

5.5.1 Assets

The survey collected data on consumer durable assets, productive assets, livestock and land. We used principal components analysis to construct an index of total asset ownership, as well as separate asset indices for consumer durables, productive assets and livestock. The asset indices are summarized in Table 5.5.1. These indices are constructed so that their mean in the data will be near zero. This makes them more useful for relative comparisons, as between regions. Total asset ownership is much higher in Amhara, and this advantage is driven by productive assets. Households in Amhara are much more likely to have a plow yoke, for example, (60% do that), than those in Oromia (40%). Same for plow beam, plow lever, plow metal support, leather tie for plow (miran), and plow blade Livestock holdings are somewhat higher in Oromia. Average land area in the sample is just under one hectare, but average landholdings are twice as large in Oromia (1.46 Ha) as in Amhara (0.62 Ha). In order to further investigate the distribution of assets by region, we present the share of the sample in each region in quartiles of the total asset index. In Oromia, 71.2 percent of households are in the first two quartiles, below median asset holdings, whereas only 34.7 percent of households in Amhara have below median asset holdings, further confirming the relative poverty of the Oromia sample.

5.5.2 Financial inclusion: access to savings, credit and financial institutions

Table 5.5.2 reports use of financial institutions in the sample. Overall, use of many financial services is low, with 12.5 of female respondents and 20.5 percent of male respondents belonging to a RUSACCO, and 10.6 percent (female) and 14.1 percent (male) belonging to a VSLA, for example. The institution with the broadest participation is the iddir, an informal insurance arrangement organized in part for funeral expenses. Participation in an iddir is at 54.0 percent for women and 60.9 percent for men and is similar across regions.

Access to credit is relatively low, as shown in Table 5.5.3, and is only modestly higher for men than for women. Only 7.5 percent of women and 11.7 percent of men took out a loan for productive purposes in the past 12 months. Similarly, only 5.8 percent of women and 9.7 percent of men took out a loan for consumption purposes in the past 12 months.

5.5.3 Aspirations

Tables 5.5.4 and 5.5.5 present responses from female and male respondents, respectively, on three education measures about their oldest child: the child's current education level, the education level that they aspire for the child to achieve, and the education level that they expect the child to achieve. In terms of current education, 36.6 percent of oldest children have no formal education and 58.1 percent have between 1st-8th grade education. Education aspirations were much higher than current achievement, which may reflect the age of the child as well. For

example, 26.8 percent of women and 26.2 percent of men aspire for their oldest child to obtain a 9th-12th grade education. Also, 12.1 percent of women and 11.8 percent of men aspire for their oldest child to obtain a college education. However, only 8.0 percent of women and 8.6 percent of men expect their child to attend college.

5.5.4 Agricultural production

Tables 5.5.6 and 5.5.7 summarize data on agricultural production in the Mehr and Belg seasons, respectively. Seventy-eight percent of households reported growing crops in the Mehr season. Households reported growing two crops on average in the main Mehr season. The most common crops grown were sorghum (in both regions), wheat and teff (in Amhara), and maize and chat (in Oromia). Total cultivated area was 12 hectares, which may reflect intercropping. Unfortunately, an error in the CAPI program used to record data during the interviews means that the total quantity and value of production were not captured for the Mehr season only, so it is not possible to present data on yields and value of production in Mehr seasons.

Farming activity was much lower in the Belg season, with only 17 percent of households growing crops. On average, 1.4 crops were grown in the Belg season. Crops grown in the Belg season were similar to the Mehr season, except that barley was an important Belg crop in Amhara and chat became even more popular as a crop in the Belg season in Oromia, with 65 percent of households that grew any Belg crops in Oromia growing chat. Revenue figures from sales of chat confirm that this is an important cash crop in Oromia.

5.5.5 Exposure to shocks

Table 5.5.8 reports male respondent recall on exposure to shocks in categories including agricultural shocks (droughts, floods, erosion, frost, pests, input and output), health shocks (death or illness) or divorce. Thirty five percent of households overall and 56 percent of households in Oromia reported a drought, the most common shock, in the last two years. No other shock was nearly as common. Frost affected 10 percent of households in Amhara, but almost no one in Oromia.

Table 5.5.1: Household Assets

	N	All	Amhara	Oromia
Household Total Asset + Land Owned Index	3,312	0.021	0.852	-1.123
		(2.852)	(2.856)	(2.419)
Consumer Durable Asset Index	3,313	0.003	0.034	-0.040
		(1.736)	(1.872)	(1.529)
Household Productive Asset Index	3,314	0.019	0.782	-1.031
		(2.728)	(2.717)	(2.371)
Household Livestock Asset Index	3,313	0.005	-0.020	0.039
		(1.406)	(1.462)	(1.325)
Area of land owned (hectares)	3,314	0.975	0.621	1.461
		(7.220)	(2.002)	(10.864)
Asset index: First quartile	3,312	0.250	0.208	0.308
		(0.433)	(0.406)	(0.462)
Asset index: Second quartile	3,312	0.250	0.139	0.403
		(0.433)	(0.346)	(0.491)
Asset index: Third quartile	3,312	0.250	0.278	0.212
		(0.433)	(0.448)	(0.409)
Asset index: Fourth quartile	3,312	0.250	0.376	0.077
		(0.433)	(0.484)	(0.266)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.5.2: Access to savings and financial institutions

	N	All	Amhara	Oromia
Primary female belongs to a RUSACCO	3,314	0.125	0.178	0.052
		(0.330)	(0.382)	(0.221)
Primary female belongs to a Village Savings and Lending Association (VSLA)	3,314	0.106	0.122	0.083
		(0.308)	(0.328)	(0.276)
Primary female belongs to a Microfinance Institution (MFI)	3,314	0.056	0.054	0.060
		(0.230)	(0.225)	(0.237)
Primary female has a bank account	3,314	0.043	0.053	0.029
		(0.203)	(0.223)	(0.169)
Primary male belongs to a RUSACCO	2,747	0.203	0.318	0.072
		(0.402)	(0.466)	(0.258)
Primary male belongs to a Village Savings and Lending Association (VSLA)	2,747	0.141	0.189	0.086
		(0.348)	(0.392)	(0.281)
Primary male belongs to a Microfinance Institution (MFI)	2,747	0.095	0.111	0.078
		(0.294)	(0.314)	(0.268)
Primary male has a bank account	2,747	0.073	0.093	0.051
		(0.260)	(0.291)	(0.219)
Primary female is a member of an Eqqub	3,314	0.040	0.043	0.035
		(0.196)	(0.203)	(0.184)
Primary female is a member of an Iddir	3,314	0.540	0.524	0.561
		(0.498)	(0.500)	(0.496)
Primary male is a member of an Eqqub	2,747	0.037	0.046	0.026
		(0.188)	(0.209)	(0.160)
Primary male is a member of an Iddir	2,747	0.609	0.619	0.598
		(0.488)	(0.486)	(0.490)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.5.3: Access to credit for production and consumption purposes

	N	All	Amhara	Oromia
Primary female took out loan for productive purposes in past 12 months	3,314	0.075	0.103	0.036
		(0.263)	(0.304)	(0.186)
Total value of productive loan primary female took out	247	6,263.036	7,267.513	2,305.400
		(5,037.270)	(5,117.610)	(1,609.196)
Primary female took out loan for consumption purposes in past 12 months	3,314	0.058	0.030	0.097
		(0.234)	(0.170)	(0.296)
Total value of consumption loan primary female took out	192	1,760.417	1,933.684	1,687.259
		(1,698.309)	(2,442.593)	(1,264.758)
Primary male took out loan for productive purposes in past 12 months	2,747	0.117	0.181	0.044
		(0.321)	(0.385)	(0.204)
Total value of productive loan primary male took out	321	7,871.589	8,456.128	5,105.464
		(11,889.591)	(9,376.952)	(19,773.328)
Primary male took out loan for consumption purposes in past 12 months	2,747	0.097	0.033	0.170
		(0.296)	(0.178)	(0.376)
Total value of consumption loan primary male took out	267	2,106.000	2,308.438	2,061.630
		(1,862.759)	(2,344.856)	(1,742.829)
Primary female had no access to loans in the past 12 months	3,067	0.129	0.048	0.233
		(0.335)	(0.213)	(0.423)
Primary female received loan from Rusacco	247	0.466	0.508	0.300
		(0.500)	(0.501)	(0.463)
Reason for loan - to buy livestock	247	0.656	0.716	0.420
		(0.476)	(0.452)	(0.499)
Total outstanding loan amount the primary female still owes	246	4,101.6	4,763.0	1,508.8
		(4,574.1)	(4,811.1)	(1,971.8)
Primary female had any difficulty in loan repayment	247	0.421	0.386	0.560
		(0.495)	(0.488)	(0.501)
Number of months primary female took out a loan for consumption purposes	192	1.458	1.439	1.467
		(0.758)	(0.802)	(0.741)

Primary female took out consumption loan for food expenditure	192	0.859	0.737	0.911
		(0.349)	(0.444)	(0.286)
Total outstanding consumption loan amount the primary female still owes	192	869.714	1,193.421	733.037
		(2,397.375)	(4,156.118)	(953.995)
Primary male had no access to loans in the past 12 months	2,426	0.156	0.056	0.253
		(0.363)	(0.230)	(0.435)
Primary male received loan from Rusacco	879	0.177	0.193	0.108
		(0.382)	(0.395)	(0.311)
Reason for loan - to buy livestock	321	0.754	0.800	0.536
		(0.431)	(0.401)	(0.503)
Total outstanding loan amount the primary male still owes	321	5,718.558	6,100.592	3,910.714
		(14,199.599)	(12,674.288)	(19,950.453)
Primary male had any difficulty in loan repayment	321	0.483	0.457	0.607
		(0.500)	(0.499)	(0.493)
Number of months primary male took out a loan for consumption purposes	267	1.667	1.354	1.735
		(1.191)	(0.699)	(1.265)
Primary male took out consumption loan for food expenditure	267	0.888	0.708	0.927
		(0.316)	(0.459)	(0.261)
Total outstanding consumption loan amount the primary male still owes	267	912.884	1,187.500	852.694
		(1,643.499)	(2,279.686)	(1,468.036)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.5.4: Present education and educational aspirations for oldest child - Responses from primary female

	N	All	Amhara	Oromia
Current education level of oldest child				
No formal education	2,619	0.366	0.373	0.359
		(0.482)	(0.484)	(0.480)
1st-8th Grade	2,619	0.581	0.551	0.613
		(0.494)	(0.498)	(0.487)
9th-12th Grade	2,619	0.044	0.062	0.024
		(0.204)	(0.241)	(0.153)
Technical/Diploma/Certificate	2,619	0.003	0.004	0.002
		(0.052)	(0.061)	(0.040)
College/University	2,619	0.001	0.002	0.000
		(0.034)	(0.047)	(0.000)
Literacy Program	2,619	0.006	0.009	0.002
		(0.075)	(0.094)	(0.049)
Aspiration education level for oldest child				
No formal education aspirations	2,619	0.068	0.075	0.061
		(0.252)	(0.263)	(0.240)
1st-8th Grade	2,619	0.197	0.233	0.158
		(0.398)	(0.423)	(0.365)
9th-12th Grade	2,619	0.268	0.256	0.281
		(0.443)	(0.436)	(0.450)
Technical/Diploma/Certificate	2,619	0.040	0.029	0.052
		(0.196)	(0.169)	(0.222)
College/University	2,619	0.121	0.086	0.159
		(0.327)	(0.281)	(0.366)
Literacy Program	2,619	0.305	0.321	0.289
		(0.461)	(0.467)	(0.453)
Expected education level of oldest child				
No formal education aspirations	2,603	0.046	0.051	0.040
		(0.209)	(0.220)	(0.196)
1st-8th Grade	2,603	0.235	0.264	0.203
		(0.424)	(0.441)	(0.402)
9th-12th Grade	2,603	0.421	0.356	0.491
		(0.494)	(0.479)	(0.500)
Technical/Diploma/Certificate	2,603	0.043	0.040	0.046
		(0.203)	(0.196)	(0.210)
College/University	2,603	0.080	0.076	0.085
		(0.272)	(0.264)	(0.280)
Literacy Program	2,603	0.175	0.213	0.135
		(0.380)	(0.409)	(0.342)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.5.5: Present education and educational aspirations of oldest child - Responses from primary male

	N	All	Amhara	Oromia
Current education level of oldest child				
No formal education	2,614	0.360	0.364	0.356
		(0.480)	(0.481)	(0.479)
1st-8th Grade	2,614	0.581	0.556	0.609
		(0.493)	(0.497)	(0.488)
9th-12th Grade	2,614	0.047	0.065	0.028
		(0.212)	(0.246)	(0.165)
Technical/Diploma/Certificate	2,614	0.005	0.009	0.002
		(0.073)	(0.094)	(0.040)
College/University	2,614	0.001	0.001	0.001
		(0.028)	(0.027)	(0.028)
Literacy Program	2,614	0.005	0.006	0.005
		(0.073)	(0.077)	(0.069)
Aspiration education level for oldest child				
No formal education aspirations	2,630	0.067	0.070	0.064
		(0.250)	(0.255)	(0.244)
1st-8th Grade	2,630	0.175	0.211	0.135
		(0.380)	(0.408)	(0.342)
9th-12th Grade	2,630	0.262	0.245	0.281
		(0.440)	(0.430)	(0.450)
Technical/Diploma/Certificate	2,630	0.037	0.027	0.049
		(0.189)	(0.162)	(0.215)
College/University	2,630	0.118	0.084	0.156
		(0.323)	(0.278)	(0.363)
Literacy Program	2,630	0.341	0.363	0.317
		(0.474)	(0.481)	(0.465)
Expected education level of oldest child				
No formal education aspirations	2,627	0.045	0.055	0.035
		(0.208)	(0.228)	(0.184)
1st-8th Grade	2,627	0.207	0.241	0.169
		(0.405)	(0.428)	(0.375)
9th-12th Grade	2,627	0.430	0.354	0.513
		(0.495)	(0.478)	(0.500)
Technical/Diploma/Certificate	2,627	0.031	0.031	0.030
		(0.173)	(0.174)	(0.171)
College/University	2,627	0.086	0.073	0.099
		(0.280)	(0.260)	(0.299)
Literacy Program	2,627	0.202	0.246	0.153
		(0.401)	(0.431)	(0.361)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.5.6: Agricultural production in Mehr 2009

	N	All	Amhara	Oromia
Number of crops cultivated in Mehr 2010	2,577	2.069	2.104	2.029
		(1.076)	(1.120)	(1.025)
Farmer grows sorghum in Mehr 2009	2,576	0.447	0.361	0.542
		(0.497)	(0.481)	(0.498)
Farmer grows wheat in Mehr 2009	2,576	0.234	0.441	0.002
		(0.423)	(0.497)	(0.050)
Farmer grows teff in Mehr 2009	2,576	0.174	0.310	0.023
		(0.379)	(0.463)	(0.150)
Farmer grows maize in Mehr 2009	2,576	0.216	0.063	0.386
		(0.411)	(0.244)	(0.487)
Farmer grows chat in Mehr 2009	2,576	0.137	0.000	0.289
		(0.344)	(0.000)	(0.454)
Area of sorghum cultivated in Mehr 2010 (hectares)	1,218	3.504	5.699	1.873
		(73.001)	(110.050)	(17.256)
Area of wheat cultivated in Mehr 2010 (hectares)	617	5.388	5.388	
		(102.084)	(102.084)	
Area of teff cultivated in Mehr 2010 (hectares)	429	0.988	0.988	
		(12.054)	(12.054)	
Area of maize cultivated in Mehr 2010(hectares)	528	0.899		0.899
		(4.586)		(4.586)
Area of chat cultivated in Mehr 2010(hectares)	443	0.543		0.543
		(1.256)		(1.256)
Total area cultivated in Mehr 2010 (hectares)	2,577	12.036	19.663	3.525
		(173.707)	(237.358)	(29.410)
Revenue from sorghum sales in Mehr 2009 (‘000 Birr)	1,151	2.144	1.107	2.916
		(24.177)	(14.181)	(29.481)
Revenue from wheat sales in Mehr 2009 (‘000 Birr)	599	1.070	1.070	
		(10.096)	(10.096)	
Revenue from teff sales in Mehr 2009 (‘000 Birr)	421	0.294	0.294	
		(1.732)	(1.732)	
Revenue from maize sales in Mehr 2009 (‘000 Birr)	470	3.342		3.342
		(48.351)		(48.351)

Revenue from chat sales in Mehr 2009 (‘000 Birr)	352	87.8		87.8
		(1,334.1)		(1,334.1)
Total revenue from crop sales in Mehr 2009 (‘000 Birr)	2,576	102.3	3.609	212.5
		(2,638.4)	(37.076)	(3,836.2)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Revenues are reported in thousands of Birr.

Table 5.5.7: Agricultural production in Belg 2009

	N	All	Amhara	Oromia
Number of crops cultivated in Belg 2009	571	1.424	1.474	1.394
		(0.685)	(0.732)	(0.655)
Farmer grows sorghum in Belg 2009	562	0.302	0.550	0.154
		(0.460)	(0.499)	(0.361)
Farmer grows barley in Belg 2009	562	0.155	0.379	0.020
		(0.362)	(0.486)	(0.140)
Farmer grows teff in Belg 2009	562	0.046	0.118	0.003
		(0.210)	(0.324)	(0.053)
Farmer grows maize in Belg 2009	562	0.169	0.014	0.262
		(0.375)	(0.119)	(0.440)
Farmer grows chat in Belg 2009	562	0.406	0.000	0.650
		(0.491)	(0.000)	(0.478)
Yield of sorghum (‘000 birr/hectare), Belg season	167	35.935	50.692	2.368
		(103.724)	(121.688)	(2.472)
Yield of barley (‘000 birr/hectare), Belg season	76	60.419	60.419	
		(277.699)	(277.699)	
Yield of teff (‘000 birr/hectare), Belg season	25	1.095	1.095	
		(1.725)	(1.725)	
Yield of maize (‘000 birr/hectare), Belg season	88	3.210		3.210
		(3.983)		(3.983)
Yield of chat (‘000 birr/hectare), Belg season	226	183.630		183.630
		(1,257.524)		(1,257.524)
Total yield (‘000 birr/hectare), Belg season	562	91.643	46.672	118.676
		(507.700)	(199.145)	(622.418)
Total value of sorghum produced in Belg 2009 (‘000 Birr)	169	14.222	20.371	0.764
		(42.572)	(50.257)	(0.623)
Total value of barley produced in Belg 2009 (‘000 Birr)	76	28.971	28.971	
		(200.706)	(200.706)	
Total value of teff produced in Belg 2009 (‘000 Birr)	25	0.318	0.318	
		(0.467)	(0.467)	
Total value of maize produced in Belg 2009 (‘000 Birr)	90	0.841		0.841
		(1.011)		(1.011)

Total value of chat produced in Belg 2009 (‘000 Birr)	228	48.181		48.181
		(314.369)		(314.369)
Total value of production in Belg 2009 (‘000 Birr)	562	34.523	24.663	40.450
		(223.365)	(127.345)	(264.850)
Area of sorghum cultivated in Belg 2009 (hectares)	168	1.298	0.617	2.819
		(9.614)	(0.522)	(17.280)
Area of barley cultivated in Belg 2009 (hectares)	80	0.508	0.508	
		(0.618)	(0.618)	
Area of teff cultivated in Belg 2009 (hectares)	25	0.410	0.410	
		(0.332)	(0.332)	
Area of maize cultivated in Belg 2009(hectares)	91	0.697		0.697
		(2.151)		(2.151)
Area of chat cultivated in Belg 2009(hectares)	227	2.251		2.251
		(26.529)		(26.529)
Total area cultivated in Belg 2009 (hectares)	571	3.489	5.229	2.468
		(26.392)	(31.842)	(22.584)
Revenue from sorghum sales in Belg 2009 (‘000 Birr)	14	9.005	10.431	0.450
		(15.376)	(16.244)	(0.071)
Revenue from barley sales in Belg 2009 (‘000 Birr)	6	16.358	16.358	
		(28.016)	(28.016)	
Revenue from teff sales in Belg 2009 (‘000 Birr)	4	0.321	0.321	
		(0.337)	(0.337)	
Revenue from maize sales in Belg 2009 (‘000 Birr)	8	0.256		0.256
		(0.274)		(0.274)
Revenue from chat sales in Belg 2009 (‘000 Birr)	174	2,138.5		2,138.5
		(27,288.6)		(27,288.6)
Total revenue from crop sales in Belg 2009 (‘000 Birr)	562	665.9	1.551	1,065.2
		(15,186.0)	(7.651)	(19,215.0)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Values of production revenues and yields are reported in thousands of Birr. Values and yields are trimmed by 1 percent at the top and bottom of the distribution to adjust for any outliers.

Table 5.5.8: Exposure to shocks

	N	All	Amhara	Oromia
Affected by a drought in the last two years	3,314	0.355	0.208	0.557
		(0.478)	(0.406)	(0.497)
Affected by a drought in 2008	3,313	0.208	0.109	0.344
		(0.406)	(0.312)	(0.475)
Affected by a drought in 2009	3,313	0.300	0.150	0.508
		(0.458)	(0.357)	(0.500)
Affected by a drought in 2010	3,313	0.127	0.090	0.178
		(0.333)	(0.286)	(0.383)
Affected by a flood in the last two years	3,314	0.035	0.055	0.009
		(0.185)	(0.227)	(0.092)
Affected by erosion in the last two years	3,314	0.028	0.038	0.014
		(0.165)	(0.191)	(0.119)
Affected by frost in the last two years	3,314	0.062	0.102	0.008
		(0.241)	(0.302)	(0.089)
Affected by pests in the last two years	3,314	0.001	0.002	0.001
		(0.035)	(0.040)	(0.027)
Affected by inputs in the last two years	3,314	0.022	0.008	0.042
		(0.148)	(0.091)	(0.200)
Affected by outputs in the last two years	3,314	0.030	0.025	0.038
		(0.172)	(0.156)	(0.191)
Affected by death in the last two years	3,314	0.000	0.001	0.000
		(0.017)	(0.023)	(0.000)
Affected by illness in the last two years	3,314	0.001	0.001	0.000
		(0.025)	(0.032)	(0.000)
Affected by divorce in the last two years	3,314	0.024	0.033	0.010
		(0.152)	(0.180)	(0.100)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

5.6 Empowerment, intimate partner violence and mental and physical wellbeing

5.6.1 Women's and men's empowerment

Women's and men's empowerment is measured using the Project-Level Women's Empowerment in Agriculture Index (pro-WEAI).[15] The index is composed of 12 indicators that capture 3 domains of empowerment: intrinsic agency (or power within); instrumental agency (or power to); and collective agency (or power with). The baseline survey collected information on 6 of the 12 indicators that make up the Pro-WEAI index: input in productive decisions, self-efficacy, attitudes about domestic violence, visiting important locations, group membership, and respect among household members. For all but visiting important locations and group membership, we collect information on both the primary male and female of the household. The six indicators are created to reflect whether the respondent achieved adequacy in that indicator; a respondent is considered adequate in a particular indicator if she or he reaches a certain threshold. The thresholds for the six indicators are described below[16]:

Input in productive decisions: Respondent either makes the decision or has at least some input into the decision, or feels he could make the decisions to at least a medium extent if he wanted to for all agricultural activities he/she participates in.

Self efficacy: respondent scores at least a 32 out of 40 on the New Generalized Self-Efficacy test.

Attitudes about domestic violence: Respondent believes husband is never justified in hitting or beating his wife.

Visiting important locations: Respondent visits at least two locations at least once per week of city, market, and family/relative; or visits at least one location at least once per month of a health facility and public meeting.

Group membership: Respondent is an active member of at least one group.

Respect among household members: Respondent respects spouse, spouse respects respondent, respondent trusts spouse, and respondent is comfortable disagreeing with spouse most of the time.

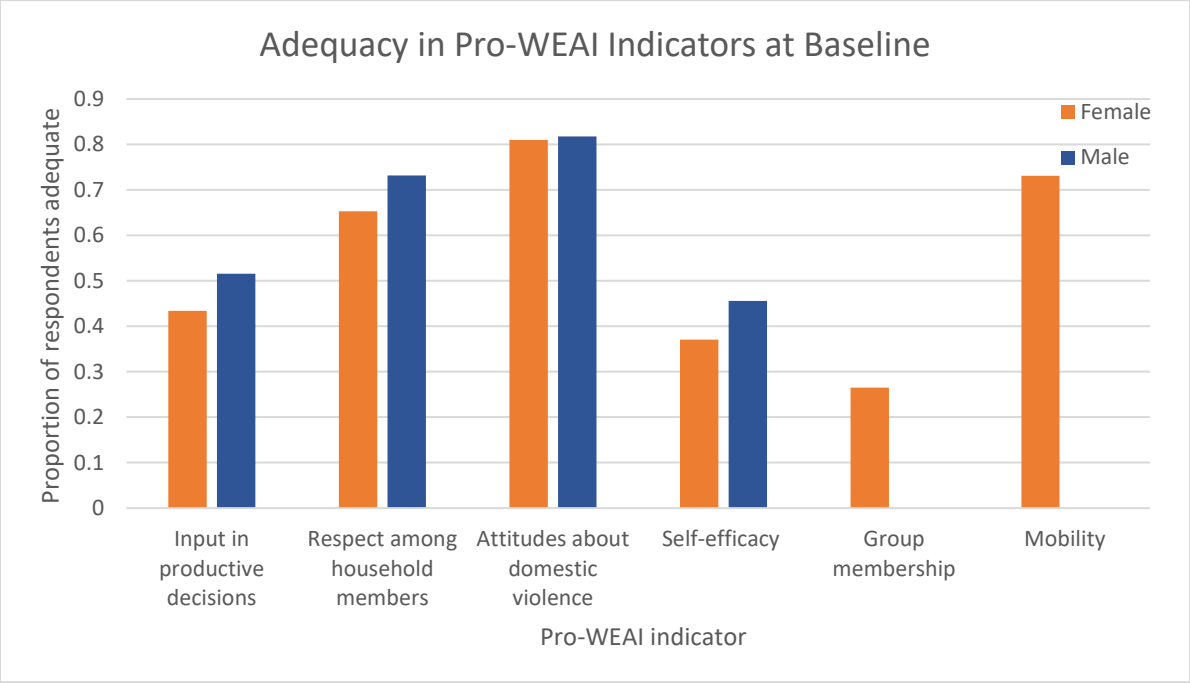
Figure 5.6.1 reveals the proportion of primary males and females in our sample that achieved adequacy across the 6 indicators. For two indicators, group membership and visiting important locations, we only collected information on the primary female. Across the four indicators where information was collected on both, men achieve higher rates of adequacy than females, however, there is virtually no difference in attitudes about domestic violence. Both men and women achieve the highest rates of adequacy in their attitudes about domestic violence and the lowest

[15] For more information on the Pro-WEAI, visit the website https://weai.ifpri.info/2018/04/27/introducing-pro-weai-a-tool-for-measuring-womens-empowerment-in-agricultural-development-projects/

[16] Cutoffs for the Pro-WEAI indicators are still not finalized, and thus may change in the future.

rates of adequacy in the self-efficacy test and group membership (female reports only). Across regions (Table 5.6.1), women in Amhara are more empowered across 4 of the 6 indicators, achieving higher rates of adequacy in input in productive decisions, group membership, respect among household members, and self-efficacy. Women in Oromia achieve higher rates of adequacy in visiting important locations and attitudes about domestic violence. Men in Amhara also have higher rates of empowerment across 3 of the 4 indicators: input into productive decisions, respect among household members, and self-efficacy.

Figure 5.4.1: Adequacy in Pro-WEAI indicators for primary male and female respondent



5.6.2 Intimate partner violence

Intimate partner violence (IPV) is a major global public health problem with multiple malign consequences for women's physical and mental health (Ellsberg et al. 2008; Kapiga et al. 2017) and is the leading cause of women's death by homicide (Devries et al. 2013). Adverse effects are transmitted intergenerationally, with IPV linked to poorer child development, nutrition, and health outcomes, as well as a greater likelihood of children also entering into abusive relationships (Aizer 2010; Fulu et al. 2017; Hasselmann and Reichenheim 2006; Karamagi et al. 2007; Koenen et al. 2003; Pollak 2004; Yount et al. 2011). Using data from 141 studies from 81 countries, Devries et al. (2013) estimate that 30 percent of all adult women have experienced physical or sexual IPV in their lifetimes. According to the 2016 EDHS, 34 percent of married women aged 15-49 in Ethiopia had experienced physical, sexual, or emotional violence from spouse (CSA and ICF 2016).

Indicators of internationally validated standardized IPV measures from the WHO Violence Against Women Instrument were administered in the baseline survey and included three types of violence (physical, sexual, emotional). This instrument was also used in the 2016 Ethiopian DHS. Violence indicators were collected in accordance with the WHO protocol on ethical guidelines for conducting research on IPV (WHO 2016). Only the primary female was interviewed for the IPV module and women who reported any violence were given the option to be referred to the Women's Affairs Committee in their woreda. Similar to WHO norms, we restrict women in our analysis to be 15-49 years old, married and alone at the time of the interview, when IPV is more prevalent, yielding a sample of 1941 women.[17]

For the three types of violence, multiple behaviorally specific questions were administered in order to reduce under-reporting. We asked if the women had ever experienced the act of violence and if she had experienced it in the last 13 months. The three types of violence were defined as follows:

> **Emotional spousal violence**: husband/partner said or did something to humiliate you in front of others; threatened to hurt or harm you or someone close to you; insulted you or made you feel bad about yourself
>
> **Physical spousal violence**: Husband/partner pushed you, shook you, or threw something at you; slapped you; twisted your arm or pulled your hair; punched you with his fist or with something that could hurt you; kicked you, dragged you, or beat you up; tried to choke you or burn you on purpose; or threatened or attack you with a knife, gun, or any other weapon
>
> **Sexual spousal violence**: Husband/partner physically forced you to have sexual intercourse with him even when you did not want to; physically forced you to perform any other sexual acts you did not want to; forced you with threats or in any other way to perform sexual acts you did not want to

In addition to the three types of violence, we administer questions on **marital control** as defined as husband/partner demonstrating at least one of the following controlling behaviours: is jealous or angry if she talks to other men; frequently accuses her of being unfaithful; does not permit her to meet her female friends; tries to limit her contact with her family; and insists on knowing where she is at all times.

Table 5.6.2 reveals lifetime and 13-month IPV rates for women 15-49 years old in the sample. Lifetime rates of violence are 14.9 percent for physical violence, 21.7 percent for emotional violence, and 5.6 percent for sexual violence. Rates of violence in the last 13 months are lower at 7.2 percent for physical violence, 12.7 percent for emotional violence, and 3.8 percent for sexual violence. Women in Oromia experience higher rates of IPV than women in Amhara for all

[17] In the protocol for the IPV module, the IPV interview should only take place when the woman is able to be alone with the enumerator for the interview. Due to an error in the CAPI program, this requirement was not enforced through skip patterns on the tablet, so some women were interviewed for the IPV module with another household member present. We only use the data from those interviews where women were able to be alone for the IPV module.

violence indicators. Rates of IPV, both lifetime and 13-month, are lower than the rates reported in the EDHS which are 24 percent for lifetime physical violence and 17 percent for physical violence in the last 12 months (CSA and ICF 2016). By region according to the EDHS, Oromia has the highest rates of lifetime physical violence at 38 percent, followed by Harari at 37 percent and Amhara at 35 percent. The low rates of IPV in our data maybe due to differences in the target population as well as differences in self-reporting bias. In future work, we will aim to measure the self-reporting bias in our sample.

5.6.3 Depression

Maternal depression is a risk factor for undernutrition as well as for delayed cognitive development in many low- and middle-income countries.[18] Both antenatal and post-natal depression can influence a child, although the pathways are different. Maternal depression may affect child outcomes from very early during pregnancy (through altered placental function, epigenetic changes, and stress reactivity) to postnatal period, infancy and childhood (via altered mother – child interactions, less affection and responsiveness, poor psychosocial stimulation, inadequate feeding, poor hygiene and health-seeking practices).[19]

The PHQ-9 - a 9-item depression diagnostic instrument - was used to assess depression in the study sites. The module asks respondents whether they experienced a set of depressive symptoms in the past week and to indicate the frequency that they experienced these symptoms, rating these on a scale of 0-3. There are various cutoffs suggested in the literature. One that is commonly used defines having mild depression severity as reporting a symptoms score between as 5 and 9; moderate severity is between 10 and 14, moderately severe between 15-19 and an individual with a score 20 and above is deemed as having symptoms of severe depression.[20]

As indicated in Table 5.6.3, very few women report symptoms of severe depression (around 1% of the total); half the sample reports no symptoms at all. The share of the sample with moderately severe or severe symptoms is only 2.4 percent. Further, 6.7 percent would report symptoms in keeping with a cutoff for mild or severe depression. This is much lower than found in other studies in low- and middle-income countries. For example, a systematic review using studies from 20 countries, including 5 from Africa indicates the prevalence of antenatal depression was 25 percent based on 51 studies using various instruments for assessing risks; 19 percent reported symptoms consistent with post-partum depression.[21] It is not clear at this time why the results from the baseline differ from these averages.

---

[18] Britto, P. et al. . "Nurturing care: promoting early childhood development." Lancet **389**(10064): 91-102. 2017.

[19] Herba, C. et al. "Maternal depression and mental health in early childhood: an examination of underlying mechanisms in low-income and middle-income countries." Lancet Psychiatry **3**(10): 983-992. 2016.

[20] Kroenke, Kurt, Robert L. Spitzer, and Janet BW Williams. "The PHQ-9: validity of a brief depression severity measure." Journal of General Internal Medicine 16, no. 9: 606-613. 2001

[21] Gelaye, B., Rondon, M.B., Araya, R. and Williams, M.A. Epidemiology of Maternal Depression, Risk Factors, and Child Outcomes in Low-income and Middle-income Countries. The Lancet Psychiatry, *3*(10), pp.973-982. 2016.

There is little difference by gender in depression rates (see male depression rates in Table 5.6.4) or in reported symptoms. The correlation between the depression severity scores of the primary female and primary male respondent is 0.7062. After regressing each of these severity scores on the enumerator ID, the correlation of residuals is 0.6948. The correlation of residuals after adding controls on household size, number of children under the age of 5, and the wealth index is 0.6194. The correlation on residuals falls to 0.6161 after adding further individual controls (age of respondent and whether the respondent has completed any formal education).

5.6.4 Health History and Child Care

Maternal Health History: Table 5.6.5 indicates utilization of ante- and post- natal services. Overall 74% had at least one antenatal visit and the majority of women received counseling on nutrition, including breastfeeding during pregnancy. Counseling was more prevalent in Amhara as was the share of women taking iron and folic acid. Similarly, nearly twice as many women in Amhara delivered in a medical facility compared to those in Oromia, with the overall average being 30 percent. Twenty percent of women received vitamin A after delivery.[22]

Child Health History: As indicated in Table 5.6.6, there is relatively little difference in the modest utilization of health facilities for child care between the two provinces. In both Amhara and Oromia, slightly more than a quarter of the children received vitamin A and similar numbers participated in growth monitoring. A slightly higher percentage of children had their nutrition status assessed using mid upper arm circumference than in terms of height or weight with very few of those who were found to be severely malnourished (11%) referred for treatment. In partial contrast, nearly half of the children who had diarrhea received oral rehydration salts (ORS), with a slightly higher share of children in Oromia receiving this treatment. Fewer – less than 20 percent overall – received zinc treatment for diarrhea.

Childcare: As shown in Table 5.6.7, very few caregivers of either gender read to children or tell them stories. However, 30 percent of women in Oromia sang to their child; twice as many as did so in Amhara. Similarly, while overall only 15 percent of women counted or drew things with the index child, this share was twice as high in Oromia compared to Amhara. Fewer men in either province sang or drew things with their child, with fewer regional differences, except for drawing and counting where participation by men was higher in Oromia. Women were far more likely to play with a child or prepare meals or bath the child, than were men. There was also a gender pattern in feeding and care for sick children but this disparity in female engagement was less than for food preparation or bathing; over 40 percent of men reported caring for a child when sick and 35 percent physically fed a child compared to 8 percent preparing food and 10 percent bathing a child. Note the recall period for caring for a sick child does not coincide with the reported illness in Table 5.6.6.

[22] This is not, however, a practice that is currently recommended by the WHO. See: https://www.who.int/elena/titles/full_recommendations/vitamina_supp/en/

Table 5.6.1: Primary female and male empowerment (Percent achieving adequacy)

	N	All	Amhara	Oromia
Primary female input in productive decisions	1,657	0.434	0.452	0.415
		(0.496)	(0.498)	(0.493)
Primary female group membership	2,867	0.265	0.294	0.217
		(0.441)	(0.456)	(0.413)
Primary female visiting important locations	3,096	0.731	0.722	0.745
		(0.443)	(0.448)	(0.436)
Primary female respect among household members	2,593	0.653	0.705	0.584
		(0.476)	(0.456)	(0.493)
Primary female attitudes about domestic violence	2,593	0.810	0.778	0.853
		(0.392)	(0.416)	(0.354)
Primary female achieved self-efficacy	3,096	0.371	0.403	0.323
		(0.483)	(0.491)	(0.468)
Primary male input in productive decisions	2,109	0.516	0.530	0.501
		(0.500)	(0.499)	(0.500)
Primary male respect among household members	2,487	0.732	0.781	0.669
		(0.443)	(0.414)	(0.471)
Primary male attitudes about domestic violence	2,487	0.818	0.811	0.827
		(0.386)	(0.392)	(0.379)
Primary male achieved self-efficacy	2,635	0.456	0.509	0.388
		(0.498)	(0.500)	(0.487)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Indicators reveal proportion of primary male or females achieving adequacy in certain domains. If a primary male or female reports that household did not participate in any of the 14 productive domains, then the indicator for input in productive decisions is coded as missing. Respect among household members and attitudes about domestic violence are only asked to married individuals.

Table 5.6.2: Intimate Partner Violence - Women age 15-49 years

	N	All	Amhara	Oromia
Experienced emotional violence in the past 13 months	1,941	0.127	0.107	0.150
		(0.333)	(0.309)	(0.357)
Experienced physical violence in the past 13 months	1,941	0.072	0.037	0.110
		(0.258)	(0.190)	(0.312)
Experienced sexual violence in the past 13 months	1,941	0.038	0.016	0.062
		(0.190)	(0.124)	(0.241)
Lifetime emotional violence	1,941	0.217	0.184	0.255
		(0.413)	(0.387)	(0.436)
Lifetime physical violence	1,941	0.149	0.108	0.195
		(0.357)	(0.310)	(0.397)
Lifetime sexual violence	1,941	0.056	0.028	0.087
		(0.230)	(0.166)	(0.282)
Marital Control by husband/partner	1,941	0.215	0.180	0.255
		(0.411)	(0.384)	(0.436)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. IPV statistics are conducted on the sample of women in the analysis who lived with their partner in the past 13 months and who were alone or with a child< 5 years during the interview.

Table 5.6.3: Maternal Depression

	N	All	Amhara	Oromia
Severity score - female	3,235	2.875 (4.232)	2.380 (4.261)	3.579 (4.089)
No depression	3,235	0.477 (0.500)	0.548 (0.498)	0.376 (0.485)
Minimal depression	3,235	0.280 (0.449)	0.270 (0.444)	0.294 (0.456)
Mild depression	3,235	0.172 (0.378)	0.123 (0.329)	0.242 (0.429)
Moderate depression	3,235	0.045 (0.208)	0.030 (0.171)	0.067 (0.251)
Moderately severe depression	3,235	0.015 (0.123)	0.014 (0.116)	0.018 (0.133)
Severe depression	3,235	0.010 (0.097)	0.015 (0.121)	0.002 (0.047)
Total number of problems felt at least several days (0-9) - female	3,235	2.291 (2.850)	1.824 (2.632)	2.954 (3.012)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Sample includes primary female respondents to the survey
who are the mother or primary caregiver of a child age 0-3.

Table 5.6.4: Primary Male Depression

	N	All	Amhara	Oromia
Severity score – male	2,680	2.804 (4.111)	1.936 (3.754)	3.820 (4.275)
No depression	2,680	0.466 (0.499)	0.583 (0.493)	0.328 (0.470)
Minimal depression	2,680	0.294 (0.456)	0.272 (0.445)	0.319 (0.466)
Mild depression	2,680	0.173 (0.378)	0.107 (0.310)	0.250 (0.433)
Moderate depression	2,680	0.046 (0.208)	0.019 (0.135)	0.077 (0.267)
Moderately severe depression	2,680	0.013 (0.114)	0.008 (0.087)	0.019 (0.138)
Severe depression	2,680	0.009 (0.094)	0.011 (0.105)	0.006 (0.080)
Total number of problems felt at least several days (0-9) - male	2,680	2.245 (2.782)	1.534 (2.399)	3.076 (2.963)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. Sample includes primary male respondents to the survey
who are the husband or partner of the primary female respondent.

Table 5.6.5: Use of antenatal and postnatal services

	N	All	Amhara	Oromia
Primary female received ANC during last pregnancy	3,253	0.760	0.814	0.685
		(0.427)	(0.389)	(0.465)
Primary female went to health facility for ANC 4+ times during last pregnancy	2,760	0.305	0.340	0.250
		(0.461)	(0.474)	(0.433)
Primary female took iron and folic acid supplements during last pregnancy	3,266	0.438	0.519	0.329
		(0.496)	(0.500)	(0.470)
Primary female received nutrition information/counseling during last pregnancy	2,999	0.552	0.597	0.495
		(0.497)	(0.491)	(0.500)
Primary female received breastfeeding information during last pregnancy	3,196	0.541	0.610	0.444
		(0.498)	(0.488)	(0.497)
Birth in a medical facility (last pregnancy)	3,077	0.306	0.382	0.214
		(0.461)	(0.486)	(0.411)
Primary female received vitamin A supplement at birth or soon after birth (last	3,077	0.201	0.214	0.186
		(0.401)	(0.410)	(0.389)
Primary female received breastfeeding help after giving birth (last pregnancy)	3,089	0.259	0.263	0.255
		(0.438)	(0.440)	(0.436)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.6.6: Child health history

	N	All	Amhara	Oromia
Index child received dose of Vitamin A in past 6 months	3,314	0.285	0.278	0.296
		(0.452)	(0.448)	(0.457)
Index child received receive any micronutrient powder in past 6 months	3,314	0.063	0.048	0.083
		(0.243)	(0.214)	(0.276)
Index child's weight was measured in past 3 months	3,314	0.274	0.258	0.296
		(0.446)	(0.438)	(0.456)
Index child's height was measured in past 3 months	3,314	0.247	0.222	0.280
		(0.431)	(0.416)	(0.449)
Index child's MUAC was measured in past 3 months	3,314	0.294	0.289	0.301
		(0.456)	(0.453)	(0.459)
Child feeding information was given at the time of measurement	1,038	0.701	0.714	0.685
		(0.458)	(0.452)	(0.465)
Index child identified as severely malnourished in past 6 months	2,913	0.094	0.095	0.093
		(0.292)	(0.293)	(0.290)
Received any referral to a facility to receive treatment for severe malnutrition	274	0.124	0.093	0.163
		(0.330)	(0.291)	(0.371)
Index child had fever in past 2 weeks	3,094	0.123	0.122	0.125
		(0.329)	(0.327)	(0.331)
Index child had cough/cold in past 2 weeks	3,094	0.148	0.145	0.151
		(0.355)	(0.353)	(0.358)
Index child had fast breathing/shortness of breath in past 2 weeks	3,087	0.032	0.029	0.036
		(0.176)	(0.168)	(0.186)
Index child had diarrhea in past 2 weeks	3,077	0.084	0.087	0.082
		(0.278)	(0.282)	(0.274)
Index child received ORS when she/he had diarrhea	255	0.482	0.455	0.518
		(0.501)	(0.500)	(0.502)
Index child received zinc to treat the diarrhea	231	0.190	0.148	0.243
		(0.394)	(0.357)	(0.431)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

Table 5.6.7: Childcare activities

	N	All	Amhara	Oromia
Read books or looked at picture books with index child	3,314	0.019	0.014	0.025
		(0.136)	(0.118)	(0.157)
Told stories index child	3,314	0.062	0.042	0.090
		(0.242)	(0.201)	(0.287)
Sang songs to or with index child	3,314	0.212	0.146	0.303
		(0.409)	(0.354)	(0.460)
Took index child outside the home	3,314	0.607	0.608	0.606
		(0.488)	(0.488)	(0.489)
Played with index child	3,314	0.786	0.779	0.796
		(0.410)	(0.415)	(0.403)
Named, counted, or drew things with or for index child	3,314	0.154	0.094	0.237
		(0.361)	(0.292)	(0.425)
Prepared food for index child who is not exclusively breastfed	3,244	0.784	0.773	0.799
		(0.412)	(0.419)	(0.401)
Physically fed index child who is not exclusively breasted	3,244	0.813	0.799	0.831
		(0.390)	(0.401)	(0.375)
Gave index child a bath	3,314	0.807	0.819	0.792
		(0.394)	(0.385)	(0.406)
Cared for the index child when they were sick	1,384	0.850	0.859	0.840
		(0.357)	(0.348)	(0.367)
Ate a meal together with index child	3,314	0.695	0.681	0.714
		(0.460)	(0.466)	(0.452)
Read books or looked at picture books with index child	2,745	0.026	0.025	0.027
		(0.159)	(0.155)	(0.163)
Told stories index child	2,745	0.060	0.044	0.079
		(0.238)	(0.205)	(0.269)
Sang songs to or with index child	2,745	0.112	0.088	0.140
		(0.316)	(0.283)	(0.347)
Took index child outside the home	2,745	0.402	0.408	0.395
		(0.490)	(0.492)	(0.489)
Played with index child	2,745	0.581	0.593	0.567
		(0.493)	(0.491)	(0.496)
Named, counted, or drew things with or for index child	2,745	0.107	0.069	0.151
		(0.310)	(0.254)	(0.358)
Prepared food for index child who is not exclusively breastfed	2,676	0.086	0.092	0.080
		(0.281)	(0.289)	(0.271)
Physically fed index child who is not exclusively breasted	2,676	0.354	0.397	0.305
		(0.478)	(0.489)	(0.461)
Gave index child a bath	2,745	0.104	0.116	0.090
		(0.305)	(0.320)	(0.287)
Cared for the index child when they were sick	1,024	0.396	0.454	0.339
		(0.489)	(0.498)	(0.474)
Ate a meal together with index child	2,745	0.431	0.433	0.430
		(0.495)	(0.496)	(0.495)

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses.

6. Balancing Tests on the Baseline Data

In this section, we report the baseline means by treatment arm of the variables presented in section 5 and conduct pairwise balancing tests of equality of means across treatment arms in the baseline survey data. The focus of these balancing tests is the four main treatment arms of the evaluation: T1: L*+N*, T2: L*+N, T3:L+N*, and C: Control. Pairwise tests across these four treatment arms leads to six balancing tests. We do not report balancing tests for the sub-randomizations on aspirations or the poultry/cash grants.

In an experimental evaluation using randomized assignment to treatment, the expected value of the difference in variable means across treatment arms is zero. In very large samples (samples with very many clusters), we may find no significant difference in means between treatment arms. However, in samples of practical size, it is possible to obtain a significant difference in means across treatment arms for some variables by chance. This is referred to as sampling error. In a simple example of sampling error, consider trying to estimate the probability of obtaining "heads" when flipping a (fair) coin. The expected value of the probability of heads is 0.5, but it is possible to get only 2 heads on 10 coin flips, yielding an estimate of 0.2. This sampling error of the estimated probability of getting heads would decline with a larger sample (more coin flips).

Thus, the interpretation of these tests of equality of means across treatment arms is not as proof that treatment assignment was random (treatment assignment was indeed random), but whether the realization of that randomization led to sampling error in some variables in the sample. It is also helpful to keep in mind, that, at a 5 percent significance level, we would expect 1 out of every 20 tests to reject equality of the means.[23] When testing a large number of variables across four treatment arms in a limited sample, it is not surprising to find some significant differences in outcomes across treatment arms.

6.1 Household demographics, child education and housing characteristics

Table 6.1.1 presents means and standard deviations by treatment arm for household demographics variables from Table 5.1.1. The table also presents the p-value from tests of the null hypothesis that the difference in means across a pair of treatment arms is statistically different from zero. Of the 66 difference-in-means tests conducted, only 2 are marginally significant at the 10 percent level, and none is statistically significant at the 5 percent level. This indicates that the data are well balanced across treatment arms on these eleven demographic variables.

[23] When testing multiple hypotheses simultaneously at a significance level of 5%, the number of significant differences that would be expected by chance is actually greater than 5%. Methods have been developed to adjust for this higher probability of finding significant effects when conducting multiple hypothesis testing, but we have decided to only present the unadjusted p-values. Readers should avoid overinterpreting significant differences that occur at rate of slightly higher than 5%.

Table 6.1.2 presents differences in means for child education variables. These are the same variables that were summarized in Table 5.1.2. The 8 variables reported leads to 48 tests of equality of means. Of these 48 tests, one is significant at the 10 percent level and one is significant at the 5 percent level. This is not more than we would expect by chance. However, net school enrollment for children age 7-13 years is significantly lower in T3 than in the Control arm. This suggests that it may improve the precision of the model used to measure treatment effects in follow-up survey rounds to control for baseline enrollment at age 7-13, for related outcomes.

Table 6.1.3 presents differences in means for housing characteristics and water sources by treatment arm. Of the 144 tests of equality of means, only one is significant at the 10 percent level. These variables are well-balanced across treatment arms in the data.

Table 6.1.1: Household demographics, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household size	5.823 (2.033)	5.862 (1.947)	5.664 (1.927)	5.749 (1.920)	0.740	0.571	0.664	0.865	0.343	0.490
Number of children under the age of 5	1.410 (0.576)	1.428 (0.557)	1.404 (0.562)	1.434 (0.571)	0.676	0.914	0.590	0.744	0.653	0.919
Female-headed household	0.218 (0.413)	0.164 (0.370)	0.212 (0.409)	0.192 (0.394)	0.536	0.456	0.593	0.166	0.159	0.879
Age of household head	39.295 (11.415)	38.760 (10.074)	38.359 (10.012)	38.256 (10.229)	0.296	0.589	0.899	0.608	0.645	0.317
Household head: Married, monogamous	0.812 (0.391)	0.854 (0.354)	0.818 (0.386)	0.834 (0.372)	0.581	0.589	0.665	0.273	0.300	0.869
Household head: Not married, divorced, widowed, separated	0.183 (0.387)	0.137 (0.344)	0.178 (0.383)	0.160 (0.366)	0.557	0.535	0.612	0.228	0.235	0.898
Household head has some education	0.280 (0.449)	0.289 (0.454)	0.288 (0.453)	0.285 (0.452)	0.876	0.919	0.926	0.800	0.989	0.801
Household head has no formal education	0.720 (0.449)	0.709 (0.454)	0.712 (0.453)	0.714 (0.452)	0.854	0.902	0.947	0.763	0.953	0.801
Household head's main activity is crop production	0.658 (0.475)	0.725 (0.447)	0.698 (0.460)	0.668 (0.471)	0.825	0.140	0.450	0.112	0.476	0.356
Age of primary female	30.619 (8.274)	30.673 (6.891)	30.496 (7.562)	30.308 (7.618)	0.495	0.388	0.669	0.911	0.704	0.804
Primary female: married, monogamous	0.847 (0.360)	0.868 (0.339)	0.822 (0.383)	0.853 (0.355)	0.876	0.650	0.397	0.521	0.170	0.472
Primary female: Not married, divorced, widowed, separated	0.152 (0.359)	0.127 (0.334)	0.178 (0.383)	0.145 (0.352)	0.846	0.608	0.361	0.452	0.132	0.446
Primary female has some education	0.190 (0.392)	0.220 (0.415)	0.200 (0.401)	0.195 (0.397)	0.851	0.417	0.855	0.333	0.538	0.719
Primary female has no formal education	0.810 (0.392)	0.780 (0.415)	0.800 (0.401)	0.805 (0.397)	0.851	0.417	0.855	0.333	0.538	0.719
Primary female's main activity is crop production	0.161 (0.367)	0.158 (0.365)	0.152 (0.359)	0.146 (0.353)	0.573	0.588	0.789	0.933	0.772	0.733
Age of primary male	38.415 (9.364)	38.100 (8.747)	38.007 (8.727)	38.045 (8.723)	0.537	0.927	0.941	0.628	0.873	0.482

Primary male: Married, monogamous	0.995	0.989	0.991	0.994	0.690	0.328	0.546	0.179	0.676	0.315
	(0.067)	(0.104)	(0.093)	(0.078)						
Primary male: Not married, divorced,	0.000	0.001	0.004	0.000	–	0.319	0.079	0.319	0.289	0.079
widowed, separated	(0.000)	(0.037)	(0.066)	(0.000)						
Primary male has some education	0.335	0.337	0.346	0.325	0.768	0.758	0.523	0.966	0.801	0.744
	(0.472)	(0.473)	(0.476)	(0.469)						
Primary male has no formal education	0.665	0.663	0.654	0.675	0.768	0.758	0.523	0.966	0.801	0.744
	(0.472)	(0.473)	(0.476)	(0.469)						
Primary male's main activity is crop	0.758	0.782	0.789	0.762	0.915	0.602	0.486	0.526	0.859	0.417
production	(0.428)	(0.413)	(0.408)	(0.426)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is from a Wald test of difference of means between the treatment arms. Where the means are identical, no p-value is reported. Standard errors are clustered at the kebele level.

Table 6.1.2: Child education, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Children age 7-18 years										
Children 7-18 years currently enrolled in school	0.675 (0.468)	0.692 (0.462)	0.649 (0.477)	0.706 (0.456)	0.436	0.680	0.093	0.689	0.246	0.538
Age at which children started school	7.635 (2.229)	7.574 (2.179)	7.670 (2.227)	7.445 (2.039)	0.332	0.456	0.223	0.753	0.599	0.863
Children who attended school at least half the time in the current school year	0.986 (0.118)	0.970 (0.170)	0.971 (0.168)	0.976 (0.154)	0.393	0.698	0.775	0.126	0.974	0.290
Number of days children attended school in the past seven days	4.693 (1.034)	4.680 (1.188)	4.626 (1.136)	4.718 (0.983)	0.797	0.686	0.302	0.899	0.582	0.502
Children age 7-13 years										
Children 7-13 years currently enrolled in school	0.681 (0.466)	0.680 (0.467)	0.632 (0.482)	0.707 (0.456)	0.527	0.443	0.038	0.983	0.228	0.280
Age at which children started school	7.076 (1.802)	7.106 (1.842)	7.150 (1.933)	7.136 (1.793)	0.697	0.841	0.929	0.846	0.785	0.656
Children who attended school at least half the time in the current school year	0.983 (0.131)	0.964 (0.187)	0.972 (0.166)	0.979 (0.144)	0.730	0.299	0.655	0.140	0.636	0.444
Number of days children attended school in the past seven days	4.696 (1.039)	4.661 (1.199)	4.640 (1.103)	4.707 (0.974)	0.908	0.631	0.456	0.749	0.843	0.593

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is from a Wald test of difference of means between the treatment arms. Standard errors are clustered at the kebele level.

Table 6.1.3: Housing characteristics and water sources, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household uses solid cooking fuels	0.993	0.995	0.996	0.995	0.615	0.936	0.631	0.616	0.782	0.397
	(0.086)	(0.068)	(0.059)	(0.071)						
Household has improved source of water - rainy season	0.623	0.624	0.678	0.576	0.520	0.487	0.120	0.989	0.394	0.419
	(0.485)	(0.485)	(0.468)	(0.495)						
Household has improved source of water - dry season	0.648	0.666	0.691	0.601	0.527	0.344	0.169	0.803	0.700	0.539
	(0.478)	(0.472)	(0.462)	(0.490)						
Time taken to fetch water										
Less than 30 mins	0.374	0.350	0.385	0.348	0.583	0.956	0.436	0.614	0.461	0.836
	(0.484)	(0.477)	(0.487)	(0.476)						
Between 30 mins-1hr	0.304	0.308	0.313	0.297	0.875	0.767	0.661	0.910	0.889	0.809
	(0.460)	(0.462)	(0.464)	(0.457)						
Between 1hr - 2hrs	0.176	0.218	0.203	0.227	0.113	0.799	0.447	0.194	0.634	0.357
	(0.381)	(0.413)	(0.402)	(0.419)						
Greater than 2 hours	0.146	0.124	0.100	0.128	0.705	0.924	0.518	0.641	0.584	0.325
	(0.354)	(0.329)	(0.300)	(0.334)						
Primary female respondent fetches the water from the source	0.838	0.828	0.819	0.849	0.663	0.444	0.277	0.743	0.752	0.518
	(0.369)	(0.377)	(0.385)	(0.358)						
Household uses the same source of drinking water for other purposes	0.755	0.770	0.753	0.789	0.401	0.634	0.333	0.680	0.607	0.956
	(0.430)	(0.421)	(0.432)	(0.408)						
Household has improved toilet	0.005	0.007	0.007	0.008	0.494	0.894	0.908	0.573	0.994	0.604
	(0.070)	(0.084)	(0.084)	(0.087)						
Toilet facility was built as part of the PSNP Public Works	0.062	0.059	0.070	0.055	0.782	0.851	0.533	0.915	0.648	0.751
	(0.241)	(0.236)	(0.255)	(0.228)						
Household has improved roof material	0.457	0.454	0.397	0.368	0.170	0.163	0.631	0.964	0.323	0.325
	(0.498)	(0.498)	(0.490)	(0.482)						
Household has improved floor material	0.057	0.053	0.099	0.054	0.934	0.968	0.269	0.897	0.238	0.281
	(0.232)	(0.223)	(0.298)	(0.226)						
Number of bedrooms	1.539	1.658	1.480	1.427	0.416	0.130	0.670	0.447	0.221	0.649
	(2.569)	(2.453)	(2.006)	(2.047)						
Household has electricity, mains	0.113	0.100	0.127	0.122	0.874	0.653	0.912	0.769	0.531	0.773
	(0.317)	(0.301)	(0.334)	(0.327)						
Household has electricity, generator	0.001	0.000	0.003	0.007	0.156	0.060	0.301	0.316	0.161	0.578

	(0.037)	(0.000)	(0.051)	(0.082)						
Household has electricity, solar panel	0.327	0.441	0.382	0.406	0.302	0.646	0.762	0.129	0.443	0.469
	(0.469)	(0.497)	(0.486)	(0.491)						
Household has electricity, other	0.053	0.041	0.059	0.069	0.542	0.255	0.696	0.561	0.445	0.831
	(0.225)	(0.198)	(0.235)	(0.254)						
Household has no electricity	0.505	0.418	0.429	0.396	0.123	0.748	0.638	0.178	0.857	0.268
	(0.500)	(0.493)	(0.495)	(0.489)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is from a Wald test of difference of means between the treatment arms. Standard errors are clustered at the kebele level.

6.2 Participation in the PSNP and start of VESA group formation

Table 6.2.1 presents balancing tests across the four treatment arms for variables on participation in the PSNP. Out of 96 tests of equality of means, one was significant at the 10 percent level and three were significant at the 5 percent level. This is still fewer rejections of equality of means than we might expect by chance. The variables significant at the 5 percent level include whether the household participated in the PSNP in Tir 2009 – Tir 2010 (for two tests) and whether the household had joined a new PSNP4 livelihood group, which was significantly more likely in T2 than in the Control group.

Table 6.2.2 reports tests of the equality of means for variables related to the initiation of SPIR activities, including participation in newly formed VESA groups and related activities. As noted in section 5, VESA groups had just begun their activities at the time of the baseline survey in fewer than six percent of communities. Table 6.2.2 shows two important trends: (i) VESA groups appear to have formed in some kebeles in the Control group (4.3%); and (ii) this early participation in VESA groups was not balanced across treatment arms, with VESA group participation significantly higher at the 5 percent in T2 than in T3 or Control. The share of households who had participated in nutrition counseling, food demonstrations or discussions about WASH through their VESA group was generally below 3 percent. This led to two tests that were significant at the 5 percent level out of 30 overall.

Table 6.2.3 tests for balancing in respondents' perceptions about the criteria for graduation from PSNP4. There is imbalance in the small share of respondents (<2% overall) that think that gaining a skill makes a household eligible for graduation (2 tests at the 10% level; 1 test at the 5% level). There was no difference across treatment arms in reasons for graduating in the last 2 years, for those that had graduated (temporarily, we presume), but there was one difference in the probability that this graduation was reported as self-graduated, with a much higher share in T1 than in T2. The difference is significant at the 5% level. Overall, 2 tests were significant at the 10% level and 2 were significant at the 5% level out of 132 tests.

Table 6.2.1: Access to the PSNP, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household has at least one member	1.000	1.000	1.000	1.000						
currently participating in PSNP	(0.000)	(0.000)	(0.000)	(0.000)						
Household participated in PW activities	0.377	0.378	0.388	0.353	0.673	0.665	0.537	0.981	0.857	0.834
between TIR 2006 and TAHISAS 2007	(0.485)	(0.485)	(0.488)	(0.478)						
Household participated in PW activities	0.602	0.611	0.555	0.560	0.449	0.367	0.938	0.872	0.335	0.411
between TIR 2007 and TAHISAS 2008	(0.490)	(0.488)	(0.497)	(0.497)						
Household participated in PW activities	0.829	0.875	0.846	0.827	0.952	0.137	0.571	0.156	0.342	0.615
between TIR 2008 and TAHISAS 2009	(0.377)	(0.331)	(0.361)	(0.379)						
Household participated in PW activities	0.906	0.956	0.914	0.877	0.402	0.008	0.272	0.055	0.104	0.785
between TIR 2009 and TAHISAS 2010	(0.292)	(0.206)	(0.280)	(0.329)						
Household head solely made the	0.650	0.605	0.620	0.657	0.870	0.188	0.358	0.250	0.696	0.450
decision about who would work on PW	(0.477)	(0.489)	(0.486)	(0.475)						
Household that sold some food	0.023	0.028	0.045	0.038	0.254	0.537	0.716	0.664	0.334	0.131
received as PW payments for cash	(0.149)	(0.166)	(0.206)	(0.192)						
Household head made the decision	0.508	0.464	0.489	0.511	0.955	0.297	0.618	0.325	0.537	0.662
about how PW transfers were used	(0.500)	(0.499)	(0.500)	(0.500)						
Household received DS payments	0.071	0.029	0.050	0.063	0.783	0.156	0.624	0.061	0.204	0.391
between TIR 2006 and TAHISAS 2007	(0.256)	(0.168)	(0.219)	(0.243)						
Household received DS payments	0.103	0.055	0.095	0.098	0.890	0.162	0.935	0.124	0.171	0.823
between TIR 2007 and TAHISAS 2008	(0.304)	(0.228)	(0.293)	(0.297)						
Household received DS payments	0.134	0.089	0.133	0.132	0.961	0.236	0.983	0.214	0.196	0.977
between TIR 2008 and TAHISAS 2009	(0.341)	(0.284)	(0.339)	(0.338)						
Household received DS payments	0.145	0.096	0.141	0.146	0.989	0.176	0.912	0.160	0.189	0.919
between TIR 2009 and TAHISAS 2010	(0.352)	(0.294)	(0.348)	(0.353)						
Households that sold some food	0.010	0.005	0.010	0.013	0.721	0.237	0.742	0.334	0.413	0.997
received as DS payments for cash	(0.102)	(0.068)	(0.102)	(0.115)						
Household head made the decision	0.399	0.394	0.384	0.426	0.672	0.595	0.457	0.932	0.851	0.788
about how DS transfers were used	(0.490)	(0.489)	(0.487)	(0.495)						
Household received government transfers NOT	0.175	0.121	0.142	0.158	0.710	0.363	0.688	0.168	0.532	0.393
related to PSNP, TIR 2009 - TAHISAS 2010	(0.380)	(0.327)	(0.349)	(0.365)						
Household has heard about new PSNP	0.745	0.728	0.750	0.694	0.223	0.378	0.182	0.684	0.595	0.904
'Livelihoods Component'	(0.436)	(0.445)	(0.433)	(0.461)						
Household has joined a new PSNP	0.398	0.474	0.380	0.351	0.348	0.019	0.557	0.143	0.075	0.731
Livelihoods Group	(0.490)	(0.500)	(0.486)	(0.478)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from a Wald test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

Table 6.2.2: Exposure to SPIR activities, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household is a member of a VESA group	0.066 (0.249)	0.091 (0.288)	0.035 (0.185)	0.043 (0.202)	0.309	0.037	0.679	0.351	0.013	0.170
Household part of a VESA group with a child under 2 years of age received counseling during a home visit	0.025 (0.156)	0.026 (0.160)	0.020 (0.139)	0.028 (0.165)	0.845	0.910	0.519	0.918	0.553	0.689
Household participated in 2 weeks of food demonstration sessions for rehabilitation of malnourished children	0.019 (0.137)	0.015 (0.120)	0.018 (0.134)	0.029 (0.169)	0.440	0.242	0.389	0.641	0.671	0.942
Household participated in VESA group discussions around WASH etc.	0.029 (0.169)	0.037 (0.189)	0.017 (0.129)	0.035 (0.185)	0.732	0.916	0.208	0.631	0.122	0.380
Household part of a VESA group participated in PW group counseling	0.037 (0.188)	0.049 (0.216)	0.030 (0.170)	0.038 (0.192)	0.934	0.571	0.584	0.517	0.249	0.648

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from a Wald test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

Table 6.2.3: Graduation from the PSNP, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household perceptions of the criteria for graduation from the PSNP										
based on income	0.687	0.696	0.681	0.692	0.888	0.915	0.779	0.818	0.723	0.885
	(0.464)	(0.460)	(0.466)	(0.462)						
based on livestock	0.105	0.093	0.125	0.082	0.352	0.642	0.108	0.624	0.224	0.454
	(0.306)	(0.290)	(0.331)	(0.274)						
based on other assets	0.082	0.080	0.100	0.095	0.608	0.479	0.832	0.922	0.378	0.495
	(0.275)	(0.272)	(0.301)	(0.293)						
months of food insecurity	0.034	0.037	0.033	0.044	0.577	0.700	0.557	0.842	0.828	0.995
	(0.180)	(0.189)	(0.180)	(0.206)						
based on skills	0.017	0.014	0.001	0.007	0.186	0.309	0.085	0.773	0.060	0.032
	(0.128)	(0.118)	(0.036)	(0.082)						
based on advice of community leader	0.010	0.011	0.008	0.008	0.744	0.608	0.945	0.807	0.555	0.676
	(0.098)	(0.106)	(0.088)	(0.089)						
self-graduation	0.004	0.004	0.005	0.012	0.320	0.306	0.418	0.919	0.787	0.838
	(0.065)	(0.062)	(0.072)	(0.109)						
none/arbitrary	0.014	0.019	0.015	0.017	0.671	0.862	0.816	0.590	0.710	0.861
	(0.117)	(0.137)	(0.123)	(0.131)						
based on other criteria	0.047	0.046	0.031	0.043	0.785	0.876	0.424	0.924	0.345	0.216
	(0.213)	(0.209)	(0.173)	(0.203)						
Household described as having graduated from the PSNP in the last 2 years	0.076	0.071	0.052	0.079	0.885	0.700	0.192	0.823	0.294	0.257
	(0.265)	(0.256)	(0.222)	(0.270)						
Among households that graduated in the last 2 years, reasons for graduating										
based on income	0.683	0.517	0.610	0.548	0.311	0.808	0.646	0.187	0.464	0.572
	(0.469)	(0.504)	(0.494)	(0.502)						
based on livestock	0.050	0.052	0.073	0.129	0.401	0.424	0.622	0.971	0.787	0.760
	(0.220)	(0.223)	(0.264)	(0.338)						
based on other assets	0.000	0.017	0.098	0.048	0.162	0.412	0.389	0.293	0.101	0.037
	(0.000)	(0.131)	(0.300)	(0.216)						
months of food insecurity	0.150	0.207	0.146	0.145	0.958	0.542	0.988	0.623	0.574	0.971
	(0.360)	(0.409)	(0.358)	(0.355)						
skills	0.067	0.052	0.024	0.065	0.979	0.834	0.468	0.842	0.528	0.550
	(0.252)	(0.223)	(0.156)	(0.248)						
advice of community leader	0.017	0.017	0.000	0.000	0.271	0.320		0.980	0.320	0.271

	(0.129)	(0.131)	(0.000)	(0.000)						
self-graduation	0.000	0.069	0.000	0.000		0.129		0.129	0.129	
	(0.000)	(0.256)	(0.000)	(0.000)						
no reason/arbitrary	0.017	0.034	0.000	0.032	0.602	0.948	0.188	0.547	0.152	0.341
	(0.129)	(0.184)	(0.000)	(0.178)						
other	0.017	0.034	0.049	0.032	0.583	0.947	0.691	0.561	0.740	0.414
	(0.129)	(0.184)	(0.218)	(0.178)						
Household described as having self-graduated from the PSNP	0.467	0.241	0.293	0.290	0.208	0.723	0.989	0.050	0.746	0.277
	(0.503)	(0.432)	(0.461)	(0.458)						
Household thought they graduated too early	0.362	0.281	0.250	0.333	0.810	0.667	0.568	0.524	0.840	0.457
	(0.485)	(0.453)	(0.439)	(0.475)						
Household stopped participating in PSNP without graduating in last 2 years	0.026	0.037	0.030	0.025	0.909	0.276	0.592	0.329	0.556	0.676
	(0.159)	(0.190)	(0.172)	(0.155)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from a Wald test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

6.3 Child nutrition and feeding practices

Table 6.3.1 presents means and standard deviations on the nutritional status indicators by treatment status and tests whether the differences in means are significantly different from zero. Of the 42 difference-in-means tests conducted, only 3 are marginally significant at the 10 percent level, and 1 is statistically significant at the 5 percent level. In particular, the proportion of children with SAM in T2 is significantly lower than the proportion in T1. Based on child anthropometrics, the randomization appears to have been very successful at selecting observably similar households across treatment groups.

Table 6.3.2 reveals that the random assignment of treatment arms was successful at balancing baseline IYCF practices. Across the 42 difference-in-means tests, only one is marginally significant at the 10 percent level.

Table 6.3.3 compares means across treatment arms for the IYCF scores of both mothers and fathers and reveals that there are no significant differences across arms. Thus, the randomization has successfully balanced nutrition knowledge characteristics across groups.

Table 6.3.1: Child anthropometry, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Height-for-age z-score (HAZ)	-1.312 (1.881)	-1.461 (1.799)	-1.410 (1.936)	-1.418 (1.865)	0.531	0.796	0.961	0.334	0.707	0.475
Proportion stunted (HAZ<-2SD)	0.355 (0.479)	0.393 (0.489)	0.383 (0.486)	0.363 (0.481)	0.831	0.441	0.564	0.309	0.770	0.394
Weight-for-height z-score (WHZ)	-0.424 (1.505)	-0.606 (1.523)	-0.613 (1.452)	-0.364 (1.494)	0.685	0.120	0.079	0.198	0.958	0.134
Proportion wasted (WHZ<-2SD)	0.124 (0.330)	0.167 (0.374)	0.144 (0.351)	0.115 (0.319)	0.745	0.118	0.266	0.222	0.484	0.486
Mid-upper arm circumference (MUAC)	13.254 (1.359)	13.364 (1.354)	13.360 (1.472)	13.364 (1.341)	0.368	0.996	0.973	0.365	0.977	0.396
Proportion with moderate acute malnutrition (11.5 cm<=MUAC<12.5 cm)	0.171 (0.377)	0.172 (0.378)	0.162 (0.369)	0.139 (0.346)	0.240	0.222	0.356	0.969	0.698	0.731
Proportion with severe acute malnutrition (MUAC<11.5)	0.085 (0.279)	0.049 (0.216)	0.059 (0.236)	0.057 (0.232)	0.071	0.572	0.895	0.014	0.461	0.078

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calcuated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

Table 6.3.2: Infant and Young Child Feeding Practices, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Children born in the last 24 months who were put to the breast within one hour	0.838 (0.368)	0.841 (0.366)	0.850 (0.358)	0.851 (0.356)	0.698	0.763	0.968	0.944	0.798	0.736
Infants 0-5 months of age who are fed exclusively breast milk	0.679 (0.469)	0.792 (0.408)	0.736 (0.443)	0.722 (0.450)	0.566	0.333	0.853	0.095	0.412	0.424
Children 12-15 months of age who are fed breast milk	1.000 (0.000)	0.990 (0.100)	0.989 (0.103)	1.000 (0.000)	–	0.313	0.317	0.313	0.959	0.316
Infants 6-8 months of age who receive solid, semi-solid or soft foods	0.431 (0.499)	0.500 (0.504)	0.431 (0.499)	0.455 (0.503)	0.799	0.610	0.802	0.465	0.470	1.000
Children 6-23 months of age who meet the minimum dietary diversity	0.012 (0.107)	0.008 (0.091)	0.021 (0.143)	0.032 (0.175)	0.134	0.073	0.552	0.646	0.376	0.523
Children 6-23 months of age who meet the minimum meal frequency	0.428 (0.495)	0.448 (0.498)	0.431 (0.496)	0.444 (0.498)	0.731	0.928	0.781	0.636	0.690	0.946
Children 6-23 months of age who receive a minimum acceptable diet	0.006 (0.076)	0.005 (0.074)	0.016 (0.124)	0.012 (0.107)	0.410	0.381	0.692	0.959	0.289	0.306

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is from the test of difference of means between the treatment arms. Standard errors are clustered at the kebele level.

Table 6.3.3: IYCF Knowledge, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Maternal IYCF knowledge score (0-14)	7.346 (2.273)	7.168 (2.159)	7.012 (2.412)	7.279 (2.165)	0.810	0.673	0.347	0.500	0.560	0.242
Maternal IYCF knowledge score (percent)	52.469 (16.236)	51.200 (15.421)	50.084 (17.227)	51.990 (15.467)	0.810	0.673	0.347	0.500	0.560	0.242
Male IYCF knowledge score (0-13)	6.411 (2.137)	6.293 (1.940)	6.271 (2.047)	6.092 (2.150)	0.121	0.293	0.361	0.540	0.909	0.483
Male IYCF knowledge score (percent)	49.316 (16.437)	48.404 (14.922)	48.242 (15.744)	46.862 (16.542)	0.121	0.293	0.361	0.540	0.909	0.483

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

6.4 Household food security, dietary diversity and consumption

Table 6.4.1 reports balancing tests for the food security and dietary diversity measures. There are no significant differences across treatment arms, so the study sample is balanced in these outcomes.

Table 6.4.2 reports differences in means across treatment arms for the consumption expenditure variables. One of the tests is significant at the 10 percent level, but out of a total of 36 comparisons.

Table 6.4.1: Food security and dietary diversity, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Food gap in months (0-12)	2.093	2.190	2.386	2.065	0.915	0.629	0.237	0.716	0.476	0.293
	(2.317)	(2.387)	(2.708)	(2.186)						
Food security index: First quartile	0.257	0.281	0.286	0.300	0.304	0.652	0.729	0.536	0.897	0.442
	(0.437)	(0.450)	(0.452)	(0.458)						
Food security index: Second quartile	0.378	0.364	0.374	0.339	0.320	0.496	0.378	0.711	0.794	0.921
	(0.485)	(0.481)	(0.484)	(0.474)						
Food security index: Third quartile	0.366	0.355	0.340	0.361	0.940	0.900	0.687	0.839	0.770	0.631
	(0.482)	(0.479)	(0.474)	(0.481)						
Number of food groups (of 12) the household consumed in the past 7 days	4.595	4.593	4.639	4.657	0.670	0.660	0.901	0.987	0.717	0.729
	(1.350)	(1.335)	(1.337)	(1.488)						
Number of food groups (of 10) women consumed the previous day or night	2.095	1.956	2.094	2.040	0.592	0.339	0.605	0.113	0.120	0.987
	(1.135)	(1.112)	(1.204)	(1.285)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.4.2: Consumption expenditure, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Consumption expenditure per month per household (Birr)	2,439 (2,124)	2,436 (2,198)	2,334 (1,765)	2,308 (1,936)	0.471	0.463	0.880	0.984	0.526	0.533
Consumption expenditure per month per adult equivalent (Birr)	599 (514)	605 (634)	596 (487)	564 (486)	0.462	0.411	0.491	0.906	0.845	0.939
Food consumption expenditure per month per adult equivalent (Birr)	479 (469)	485 (559)	466 (439)	460 (449)	0.657	0.544	0.882	0.872	0.625	0.749
Non-food consumption expenditure per month per adult equivalent (Birr)	120 (134)	120 (172)	129 (166)	104 (120)	0.179	0.311	0.054	0.962	0.550	0.494
Calories (kcal) of food consumption per adult equivalent per day	3,716 (5,140)	4,863 (19,910)	3,564 (4,618)	3,349 (3,301)	0.341	0.069	0.485	0.185	0.120	0.697
% of population living in households below the $1.90 poverty line	0.440 (0.497)	0.467 (0.499)	0.444 (0.497)	0.459 (0.499)	0.707	0.873	0.771	0.563	0.627	0.930
% of population living in households below the $1.25 poverty line	0.224 (0.417)	0.215 (0.411)	0.213 (0.410)	0.236 (0.425)	0.789	0.598	0.577	0.817	0.972	0.793

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

6.5 Livelihood outcomes: assets, financial inclusion, aspirations, agricultural production and exposure to shocks

Table 6.5.1 presents estimates of the difference in means for household assets across treatment arms, including the distribution of assets into total asset quartiles. There are no significant differences in means for the asset indices across treatment arms. However, there are differences in the share of households in each asset total asset quartile by treatment arm, with three estimates significant at the 10 percent level and two estimates significant at the 5 percent level.

For variables on access to savings and financial institutions (Table 6.5.2), there are 1 out of 72 tests significant at the 10 percent level. For variables on access to credit (Table 6.5.3), there 2 significant at the 10 percent level and 3 significant at the 5 percent level out of 132 comparisons, respectively.

Table 6.5.4 reports the education level of the oldest child as well as the female parent or caregiver's education aspirations and schooling expectations for that child. T3 has a higher share of children with no formal education (40.9%) than the other treatment arms and the difference with T2 (31.6%) is significant at the 5 percent level. Other variables report differences in schooling or aspirations at relatively narrow categories of school levels, so there are some small differences in means there. Table 6.5.5 reports the education level of the oldest child from the male respondent interview as well as the primary male respondent's education aspirations and schooling expectations for that child. Overall, four tests are significant at the 10 percent level and 4 tests are significant at the 5 percent level, out of 108 tests, which is within the frequency of significant test results that would occur by chance. The reported education level of the oldest child in the female and male interviews are quite similar, though some differences do exist which leads to different results in balancing tests. In the male interview, T2 has a lower share of children with no formal education, leading to significant differences with Control (5% level) and T3 (10% level). The other differences in education level or in education aspirations are sufficient that these baseline differences should be controlled for in future estimates of treatment effects of SPIR.

Table 6.5.6 reports balancing tests for exposure to shocks. Results are presented for 13 types of shocks. Overall the mean exposure to shocks is well balanced across treatment arms with only one test significant at the 10 percent level and two tests significant at the 5 percent level out of 78 tests. Reported recent exposure to drought (in EC 2010) was higher in T1 and T3 than in Control, which should be considered as we explore differences in agricultural and related outcomes. The rate of divorce is also higher in T2 than in Control.

Table 6.5.1: Household Assets, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Household Total Asset + Land Owned Index	-0.135 (2.906)	0.343 (2.735)	-0.048 (2.860)	-0.094 (2.892)	0.906	0.181	0.884	0.179	0.240	0.802
Consumer Durable Asset Index	-0.064 (1.602)	0.149 (1.944)	0.025 (1.731)	-0.111 (1.621)	0.784	0.217	0.485	0.308	0.590	0.643
Household Productive Asset Index	-0.131 (2.762)	0.338 (2.630)	-0.062 (2.744)	-0.084 (2.757)	0.882	0.171	0.942	0.158	0.197	0.830
Household Livestock Asset Index	-0.012 (1.498)	-0.023 (1.213)	-0.039 (1.333)	0.100 (1.568)	0.439	0.292	0.251	0.936	0.876	0.842
Area of land owned (hectares)	0.781 (2.076)	0.959 (8.614)	1.373 (10.758)	0.761 (2.942)	0.914	0.550	0.349	0.602	0.561	0.369
Asset index: First quartile	0.280 (0.449)	0.202 (0.402)	0.256 (0.436)	0.265 (0.441)	0.679	0.060	0.802	0.050	0.166	0.561
Asset index: Second quartile	0.251 (0.434)	0.242 (0.428)	0.251 (0.434)	0.257 (0.437)	0.869	0.694	0.874	0.812	0.806	0.995
Asset index: Third quartile	0.220 (0.414)	0.301 (0.459)	0.243 (0.429)	0.233 (0.423)	0.677	0.029	0.780	0.010	0.065	0.494
Asset index: Fourth quartile	0.249 (0.433)	0.255 (0.436)	0.251 (0.434)	0.245 (0.430)	0.923	0.827	0.887	0.912	0.929	0.974

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.5.2: Access to savings and financial institutions, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Primary female belongs to a RUSACCO	0.150	0.132	0.113	0.104	0.120	0.389	0.757	0.567	0.515	0.161
	(0.357)	(0.339)	(0.316)	(0.306)						
Primary female belongs to a Village Savings and Lending Association (VSLA)	0.116	0.113	0.098	0.095	0.440	0.499	0.903	0.901	0.553	0.486
	(0.321)	(0.317)	(0.298)	(0.294)						
Primary female belongs to a Microfinance Institution (MFI)	0.058	0.060	0.052	0.055	0.907	0.855	0.880	0.957	0.711	0.781
	(0.234)	(0.237)	(0.221)	(0.229)						
Primary female has a bank account	0.046	0.055	0.039	0.031	0.311	0.121	0.587	0.567	0.282	0.614
	(0.209)	(0.228)	(0.193)	(0.174)						
Primary male belongs to a RUSACCO	0.200	0.232	0.176	0.201	0.977	0.511	0.543	0.472	0.192	0.537
	(0.400)	(0.422)	(0.381)	(0.401)						
Primary male belongs to a Village Savings and Lending Association (VSLA)	0.120	0.162	0.160	0.118	0.939	0.138	0.211	0.186	0.954	0.261
	(0.326)	(0.369)	(0.367)	(0.323)						
Primary male belongs to a Microfinance Institution (MFI)	0.084	0.094	0.093	0.110	0.444	0.593	0.582	0.748	0.975	0.774
	(0.278)	(0.292)	(0.291)	(0.314)						
Primary male has a bank account	0.072	0.091	0.071	0.056	0.393	0.096	0.386	0.395	0.351	0.965
	(0.259)	(0.288)	(0.258)	(0.230)						
Primary female is a member of an Eqqub	0.045	0.037	0.042	0.035	0.551	0.868	0.659	0.622	0.743	0.889
	(0.207)	(0.190)	(0.201)	(0.184)						
Primary female is a member of an Iddir	0.523	0.558	0.558	0.518	0.944	0.533	0.555	0.584	0.997	0.603
	(0.500)	(0.497)	(0.497)	(0.500)						
Primary male is a member of an Eqqub	0.027	0.042	0.044	0.033	0.644	0.531	0.570	0.222	0.935	0.325
	(0.162)	(0.201)	(0.205)	(0.180)						
Primary male is a member of an Iddir	0.559	0.634	0.639	0.601	0.552	0.617	0.584	0.244	0.932	0.243
	(0.497)	(0.482)	(0.481)	(0.490)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.5.3: Access to credit for production and consumption purposes, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Primary female took out loan for productive purposes in past 12 mts	0.088 (0.283)	0.088 (0.283)	0.062 (0.242)	0.060 (0.238)	0.188	0.179	0.917	0.982	0.157	0.169
Total value of productive loan primary female took out (Birr)	6,755.634 (6,326.162)	6,409.733 (4,688.810)	5,653.585 (4,605.792)	5,978.125 (3,770.545)	0.461	0.680	0.769	0.741	0.491	0.319
Primary female took out loan for consumption purposes in past 12 mts	0.048 (0.215)	0.048 (0.214)	0.059 (0.235)	0.078 (0.268)	0.130	0.116	0.279	0.979	0.515	0.545
Total value of consumption loan primary female took out (Birr)	1,928.205 (1,413.227)	1,744.146 (1,451.611)	1,851.400 (2,376.624)	1,592.258 (1,346.925)	0.284	0.686	0.568	0.645	0.836	0.871
Primary male took out loan for productive purposes in past 12 mts	0.113 (0.317)	0.143 (0.350)	0.112 (0.316)	0.097 (0.296)	0.555	0.098	0.561	0.262	0.236	0.978
Total value of productive loan primary male took out (Birr)	7,685.520 (9,188.866)	7,828.343 (11,163.408)	6,943.377 (6,529.751)	9,277.344 (18,936.414)	0.556	0.602	0.368	0.934	0.562	0.591
Primary male took out loan for consumption purposes in past 12 mts	0.111 (0.315)	0.080 (0.272)	0.087 (0.283)	0.112 (0.316)	0.986	0.379	0.487	0.362	0.822	0.473
Total value of consumption loan primary male took out (Birr)	2,494.419 (1,868.562)	1,685.593 (1,240.379)	2,392.667 (2,672.276)	1,820.338 (1,308.482)	0.005	0.630	0.164	0.009	0.121	0.808
Primary female had no access to loans in the past 12 months	0.121 (0.326)	0.118 (0.322)	0.142 (0.350)	0.134 (0.340)	0.729	0.655	0.803	0.931	0.508	0.574
Primary female received loan from Rusacco	0.521 (0.503)	0.573 (0.498)	0.264 (0.445)	0.438 (0.501)	0.464	0.236	0.118	0.674	0.012	0.036
Reason for loan - to buy livestock	0.676 (0.471)	0.693 (0.464)	0.623 (0.489)	0.604 (0.494)	0.650	0.557	0.906	0.874	0.507	0.645
Total outstanding loan amount the primary female owes	4,687.514 (5,215.799)	3,785.347 (4,363.874)	3,810.340 (4,416.757)	4,062.813 (4,097.797)	0.509	0.745	0.791	0.260	0.975	0.334
Primary female had any difficulty in loan repayment	0.423 (0.497)	0.387 (0.490)	0.283 (0.455)	0.625 (0.489)	0.051	0.015	0.001	0.694	0.227	0.135
Primary female took out loan for consumption purposes in past 12 mts	0.048 (0.215)	0.048 (0.214)	0.059 (0.235)	0.078 (0.268)	0.130	0.116	0.279	0.979	0.515	0.545
Number of months primary female took out a loan for consumption	1.513 (0.721)	1.415 (0.741)	1.300 (0.707)	1.581 (0.821)	0.737	0.492	0.110	0.705	0.628	0.281
Primary female took out loan for	0.846	0.902	0.840	0.855	0.916	0.475	0.843	0.504	0.413	0.946

food expenditure	(0.366)	(0.300)	(0.370)	(0.355)						
Total outstanding consumption loan amount the primary female owes	1,444.872 (4,470.382)	682.927 (998.975)	795.500 (1,847.058)	691.290 (1,371.142)	0.313	0.975	0.761	0.301	0.714	0.394
Primary male had no access to loans in the past 12 months	0.132 (0.339)	0.159 (0.366)	0.166 (0.372)	0.166 (0.372)	0.386	0.872	0.995	0.533	0.880	0.407
Primary male received loan from Rusacco	0.183 (0.388)	0.260 (0.440)	0.116 (0.321)	0.152 (0.359)	0.552	0.084	0.486	0.184	0.011	0.138
Reason for loan - to buy livestock	0.773 (0.421)	0.771 (0.422)	0.675 (0.471)	0.797 (0.406)	0.737	0.709	0.148	0.980	0.271	0.270
Total outstanding loan amount the primary male still owes (Birr)	5,410.187 (9,013.898)	4,488.076 (11,686.821)	4,917.922 (5,477.500)	9,061.953 (25,539.577)	0.283	0.196	0.222	0.541	0.771	0.658
Primary male had any difficulty in loan repayment	0.533 (0.502)	0.524 (0.502)	0.442 (0.500)	0.406 (0.495)	0.287	0.288	0.780	0.915	0.410	0.401
Number of months primary male took out a loan for consumption	1.703 (1.003)	1.559 (0.749)	1.967 (1.948)	1.473 (0.726)	0.246	0.677	0.154	0.528	0.264	0.458
Primary male took out consumption loan for food expenditure	0.905 (0.295)	0.949 (0.222)	0.783 (0.415)	0.905 (0.295)	1.000	0.347	0.152	0.372	0.050	0.159
Total outstanding consumption loan amount the primary male still owes (Birr)	931.622 (1,643.604)	620.339 (816.327)	1,259.167 (2,446.313)	846.622 (1,265.759)	0.719	0.255	0.293	0.125	0.091	0.404

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.5.4: Present education and educational aspirations of oldest child - Responses from primary female, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Current level of education of oldest child										
No formal education	0.372	0.316	0.409	0.371	0.985	0.127	0.309	0.156	0.011	0.358
	(0.484)	(0.465)	(0.492)	(0.483)						
1st-8th Grade	0.575	0.617	0.552	0.577	0.947	0.263	0.491	0.281	0.075	0.574
	(0.495)	(0.487)	(0.498)	(0.494)						
9th-12th Grade	0.043	0.055	0.029	0.047	0.743	0.596	0.154	0.395	0.063	0.249
	(0.202)	(0.228)	(0.168)	(0.212)						
Technical/Diploma/Certificate	0.003	0.004	0.003	0.000	0.151	0.075	0.155	0.712	0.686	0.974
	(0.056)	(0.066)	(0.055)	(0.000)						
College/University	0.003	0.001	0.000	0.000	0.161	0.322		0.526	0.322	0.161
	(0.056)	(0.038)	(0.000)	(0.000)						
Literacy Program	0.005	0.006	0.008	0.005	0.994	0.834	0.488	0.839	0.746	0.494
	(0.069)	(0.076)	(0.087)	(0.068)						
Level of education you would like your oldest child to achieve										
No formal education aspirations	0.067	0.053	0.066	0.089	0.299	0.044	0.250	0.476	0.431	0.988
	(0.249)	(0.225)	(0.249)	(0.284)						
1st-8th Grade	0.198	0.201	0.191	0.198	0.993	0.933	0.863	0.940	0.794	0.857
	(0.399)	(0.401)	(0.394)	(0.399)						
9th-12th Grade	0.265	0.311	0.255	0.237	0.500	0.075	0.620	0.271	0.121	0.776
	(0.442)	(0.463)	(0.436)	(0.426)						
Technical/Diploma/Certificate	0.036	0.043	0.036	0.044	0.506	0.938	0.461	0.551	0.505	0.977
	(0.188)	(0.204)	(0.187)	(0.206)						
College/University	0.116	0.116	0.136	0.119	0.913	0.912	0.535	0.998	0.474	0.467
	(0.320)	(0.320)	(0.343)	(0.324)						
Literacy Program	0.319	0.276	0.316	0.313	0.898	0.350	0.941	0.277	0.291	0.954
	(0.466)	(0.447)	(0.465)	(0.464)						
Level of education that you think your oldest child will reach in 10 years										
No formal education aspirations	0.049	0.036	0.038	0.061	0.485	0.091	0.144	0.380	0.886	0.482
	(0.216)	(0.187)	(0.192)	(0.239)						
1st-8th Grade	0.243	0.229	0.252	0.216	0.454	0.693	0.191	0.717	0.461	0.796

	(0.429)	(0.420)	(0.434)	(0.412)						
9th-12th Grade	0.397	0.397	0.438	0.454	0.181	0.148	0.659	0.990	0.262	0.305
	(0.490)	(0.490)	(0.496)	(0.498)						
Technical/Diploma/Certificate	0.035	0.055	0.034	0.048	0.242	0.576	0.202	0.079	0.063	0.908
	(0.183)	(0.229)	(0.180)	(0.214)						
College/University	0.083	0.084	0.090	0.062	0.244	0.165	0.129	0.961	0.779	0.763
	(0.277)	(0.278)	(0.286)	(0.242)						
Literacy Program	0.194	0.198	0.149	0.158	0.280	0.269	0.752	0.908	0.148	0.143
	(0.396)	(0.399)	(0.357)	(0.365)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.5.5: Present education and educational aspirations of oldest child - Responses from primary male, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Current level of education of oldest child										
No formal education	0.362	0.313	0.385	0.386	0.536	0.040	0.980	0.194	0.053	0.568
	(0.481)	(0.464)	(0.487)	(0.487)						
1st-8th Grade	0.585	0.609	0.569	0.560	0.510	0.159	0.803	0.502	0.272	0.691
	(0.493)	(0.488)	(0.496)	(0.497)						
9th-12th Grade	0.049	0.062	0.031	0.045	0.832	0.286	0.227	0.421	0.024	0.185
	(0.215)	(0.242)	(0.173)	(0.208)						
Technical/Diploma/Certificate	0.000	0.010	0.006	0.005	0.084	0.268	0.718	0.015	0.430	0.040
	(0.000)	(0.100)	(0.078)	(0.068)						
College/University	0.000	0.001	0.000	0.002	0.317	0.956	0.317	0.323	0.323	_
	(0.000)	(0.038)	(0.000)	(0.040)						
Literacy Program	0.005	0.004	0.009	0.003	0.648	0.712	0.275	0.919	0.391	0.436
	(0.069)	(0.066)	(0.096)	(0.056)						
Level of education you would like your oldest child to achieve										
No formal education aspirations	0.072	0.054	0.066	0.076	0.868	0.210	0.603	0.368	0.485	0.771
	(0.259)	(0.227)	(0.249)	(0.265)						
1st-8th Grade	0.183	0.163	0.158	0.197	0.702	0.334	0.249	0.578	0.872	0.467
	(0.387)	(0.370)	(0.365)	(0.398)						
9th-12th Grade	0.261	0.286	0.257	0.241	0.559	0.210	0.608	0.519	0.410	0.901
	(0.440)	(0.452)	(0.437)	(0.428)						
Technical/Diploma/Certificate	0.038	0.041	0.024	0.046	0.510	0.710	0.045	0.754	0.087	0.194
	(0.191)	(0.200)	(0.153)	(0.210)						
College/University	0.109	0.123	0.125	0.116	0.786	0.782	0.723	0.549	0.940	0.484
	(0.311)	(0.329)	(0.331)	(0.320)						
Literacy Program	0.337	0.332	0.370	0.325	0.774	0.869	0.307	0.902	0.381	0.445
	(0.473)	(0.471)	(0.483)	(0.469)						
Level of education that you think your oldest child will reach in 10 years										
No formal education aspirations	0.055	0.033	0.047	0.048	0.672	0.302	0.962	0.125	0.313	0.632
	(0.228)	(0.179)	(0.211)	(0.213)						
1st-8th Grade	0.199	0.199	0.218	0.211	0.690	0.666	0.820	0.991	0.539	0.568

	(0.400)	(0.400)	(0.413)	(0.408)						
9th-12th Grade	0.401	0.426	0.427	0.465	0.096	0.315	0.303	0.501	0.969	0.448
	(0.491)	(0.495)	(0.495)	(0.499)						
Technical/Diploma/Certificate	0.036	0.034	0.029	0.024	0.194	0.327	0.597	0.888	0.614	0.464
	(0.187)	(0.182)	(0.167)	(0.153)						
College/University	0.088	0.092	0.095	0.067	0.262	0.146	0.140	0.833	0.865	0.725
	(0.283)	(0.289)	(0.294)	(0.250)						
Literacy Program	0.221	0.215	0.184	0.186	0.345	0.432	0.963	0.882	0.349	0.264
	(0.415)	(0.411)	(0.388)	(0.389)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calcuated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.5.6: Exposure to shocks, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Affected by a drought in the last two years	0.375	0.349	0.375	0.317	0.336	0.615	0.325	0.685	0.681	0.996
	(0.485)	(0.477)	(0.484)	(0.466)						
Affected by a drought in 2008	0.206	0.204	0.208	0.213	0.882	0.866	0.907	0.978	0.943	0.965
	(0.404)	(0.403)	(0.406)	(0.410)						
Affected by a drought in 2009	0.321	0.299	0.311	0.270	0.381	0.628	0.466	0.731	0.845	0.872
	(0.467)	(0.458)	(0.463)	(0.444)						
Affected by a drought in 2010	0.143	0.125	0.163	0.075	0.089	0.187	0.028	0.719	0.437	0.683
	(0.350)	(0.331)	(0.370)	(0.264)						
Affected by a flood in the last two years	0.038	0.021	0.046	0.036	0.908	0.345	0.618	0.228	0.121	0.670
	(0.192)	(0.143)	(0.209)	(0.187)						
Affected by erosion in the last two years	0.036	0.019	0.034	0.024	0.365	0.653	0.370	0.232	0.225	0.892
	(0.186)	(0.135)	(0.181)	(0.153)						
Affected by frost in the last two years	0.067	0.047	0.076	0.059	0.787	0.581	0.519	0.459	0.225	0.765
	(0.250)	(0.211)	(0.265)	(0.236)						
Affected by pests in the last two years	0.001	0.001	0.001	0.001	0.993	0.959	0.962	0.966	0.997	0.969
	(0.035)	(0.034)	(0.034)	(0.035)						
Affected by inputs in the last two years	0.010	0.029	0.023	0.026	0.186	0.868	0.834	0.155	0.699	0.143
	(0.099)	(0.168)	(0.151)	(0.160)						
Affected by outputs in the last two years	0.030	0.025	0.039	0.029	0.946	0.739	0.483	0.682	0.307	0.516
	(0.170)	(0.155)	(0.193)	(0.168)						
Affected by death in the last two years	0.000	0.000	0.001	0.000	–	–	0.318	–	0.318	0.318
	(0.000)	(0.000)	(0.034)	(0.000)						
Affected by illness in the last two years	0.000	0.001	0.001	0.000	–	0.318	0.318	0.318	0.997	0.318
	(0.000)	(0.034)	(0.034)	(0.000)						
Affected by divorce in the last two years	0.026	0.014	0.023	0.031	0.551	0.033	0.358	0.098	0.164	0.742
	(0.159)	(0.118)	(0.151)	(0.174)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calcuated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

6.6 Empowerment, intimate partner violence and mental and physical wellbeing

Table 6.6.1 reveals that the random assignment of treatment arms was successful at balancing baseline empowerment characteristics across arms. Across the 60 difference-in-means tests, only 2 are marginally significant at the 10 percent level and one at the 5 percent level.

Table 6.6.2 reveals that randomization was successful at balancing baseline IPV characteristics across treatment arms. Across 42 difference-in-means tests only one is significant at the 5 percent level and one at the 10 percent level. In particular, women in T3 experience more lifetime emotional violence than women in T1 and T2.

Table 6.6.3 shows that the scores for measures of maternal depression are balanced across all treatment arms. Table 6.6.4 shows that the depression measures for male respondents are well balanced, with one test significant at the 10 percent level and one significant at the 5 percent level.

Balancing tests for use of prenatal and antenatal services are presented in Table 6.6.5. Across 48 difference-in-means tests, one is significant at the 5 percent level and three at the 10 percent level.

The child health variables are well balanced across treatment arms (Table 6.6.6), with two tests of differences in means significant at the 5 percent level out of 84 tests.

Table 6.6.7 presents results on childcare activities. Here, there are some differences in means, with two tests significant at the 10 percent level and eight tests significant at the 5 percent level out of 132 tests. Many of these differences arise because respondents in households assigned to the T2 arm were more likely to prepare food for the young child and also were more likely to eat with the child.

Table 6.6.1: Primary female and male empowerment (Percent achieving adequacy) , by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Primary female input in productive decisions	0.421 (0.494)	0.466 (0.499)	0.434 (0.496)	0.432 (0.496)	0.836	0.521	0.978	0.424	0.547	0.818
Primary female group membership	0.316 (0.465)	0.234 (0.424)	0.259 (0.438)	0.251 (0.434)	0.165	0.689	0.845	0.062	0.517	0.197
Primary female visiting important locations	0.753 (0.431)	0.720 (0.449)	0.716 (0.451)	0.770 (0.421)	0.613	0.142	0.124	0.354	0.922	0.315
Primary female respect among household members	0.648 (0.478)	0.668 (0.471)	0.628 (0.484)	0.638 (0.481)	0.798	0.472	0.810	0.627	0.335	0.613
Primary female attitudes about domestic violence	0.824 (0.381)	0.799 (0.401)	0.789 (0.408)	0.805 (0.396)	0.595	0.865	0.669	0.493	0.805	0.335
Primary female achieved self-efficacy	0.338 (0.473)	0.373 (0.484)	0.352 (0.478)	0.410 (0.492)	0.086	0.394	0.162	0.429	0.625	0.755
Primary male input in productive decisions	0.533 (0.499)	0.540 (0.499)	0.504 (0.500)	0.521 (0.500)	0.804	0.645	0.700	0.858	0.378	0.532
Primary male respect among household members	0.739 (0.439)	0.731 (0.444)	0.688 (0.464)	0.724 (0.448)	0.671	0.850	0.375	0.814	0.276	0.179
Primary male attitudes about domestic violence	0.812 (0.391)	0.805 (0.397)	0.781 (0.414)	0.839 (0.368)	0.415	0.313	0.094	0.830	0.482	0.355
Primary male achieved self-efficacy	0.448 (0.498)	0.445 (0.497)	0.437 (0.496)	0.474 (0.500)	0.518	0.471	0.395	0.932	0.861	0.797

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calcuated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Input in productive decisions relates to primary females who make the decision has input in the decision or feels like they could make decision if wanted to about ALL of the agricultural activities they participated in.

Table 6.6.2: Attitude towards domestic violence and women's mobility, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
A primary female believes that										
a husband is justified in beating his wife if she goes out without telling him	0.112 (0.316)	0.134 (0.341)	0.157 (0.364)	0.135 (0.342)	0.432	0.982	0.494	0.441	0.478	0.153
a husband is justified in beating his wife if she neglects the children	0.072 (0.259)	0.100 (0.300)	0.117 (0.322)	0.094 (0.292)	0.371	0.832	0.387	0.269	0.516	0.078
a husband is justified in beating his wife if she argues with him	0.101 (0.301)	0.107 (0.309)	0.127 (0.334)	0.111 (0.314)	0.705	0.888	0.575	0.818	0.480	0.325
a husband is justified in beating his wife if she burns the food	0.074 (0.262)	0.082 (0.275)	0.091 (0.287)	0.072 (0.258)	0.929	0.699	0.444	0.767	0.762	0.514
it is acceptable to travel alone to the market	0.536 (0.499)	0.513 (0.500)	0.525 (0.500)	0.461 (0.499)	0.111	0.239	0.164	0.608	0.784	0.813
it is acceptable to travel alone to visit friends/family	0.568 (0.496)	0.557 (0.497)	0.560 (0.497)	0.496 (0.500)	0.116	0.153	0.143	0.787	0.933	0.854
it is acceptable to travel alone to the health center	0.552 (0.498)	0.566 (0.496)	0.553 (0.498)	0.516 (0.500)	0.454	0.236	0.421	0.727	0.738	0.978
A primary male believes that										
a husband is justified in beating his wife if she goes out without telling him	0.122 (0.328)	0.127 (0.334)	0.147 (0.354)	0.098 (0.297)	0.381	0.296	0.099	0.863	0.514	0.412
a husband is justified in beating his wife if she neglects the children	0.082 (0.274)	0.093 (0.290)	0.112 (0.316)	0.076 (0.265)	0.802	0.459	0.135	0.627	0.405	0.207
a husband is justified in beating his wife if she argues with him	0.085 (0.279)	0.104 (0.306)	0.130 (0.337)	0.096 (0.295)	0.617	0.762	0.179	0.398	0.297	0.043
a husband is justified in beating his wife if she burns the food	0.058 (0.234)	0.072 (0.259)	0.073 (0.261)	0.055 (0.229)	0.880	0.365	0.323	0.481	0.956	0.437
it is acceptable to travel alone to the market	0.515 (0.500)	0.499 (0.500)	0.526 (0.500)	0.453 (0.498)	0.194	0.327	0.131	0.706	0.526	0.792
it is acceptable to travel alone to visit friends/family	0.565 (0.496)	0.563 (0.496)	0.573 (0.495)	0.491 (0.500)	0.107	0.110	0.070	0.958	0.810	0.856
it is acceptable to travel alone to the health center	0.529 (0.500)	0.548 (0.498)	0.555 (0.497)	0.520 (0.500)	0.841	0.530	0.479	0.628	0.886	0.561

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calcuated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level.

Table 6.6.3: Maternal depression, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Severity score – female (0-27)	3.119 (4.338)	2.559 (3.942)	2.812 (4.273)	3.040 (4.361)	0.863	0.299	0.591	0.222	0.546	0.463
No depression (0)	0.433 (0.496)	0.500 (0.500)	0.507 (0.500)	0.464 (0.499)	0.509	0.439	0.316	0.178	0.885	0.109
Minimal depression (1-4)	0.297 (0.457)	0.287 (0.453)	0.261 (0.439)	0.276 (0.447)	0.447	0.652	0.580	0.701	0.290	0.194
Mild depression (5-9)	0.196 (0.397)	0.165 (0.371)	0.155 (0.362)	0.176 (0.381)	0.467	0.730	0.419	0.323	0.727	0.121
Moderate depression (10-14)	0.050 (0.218)	0.029 (0.167)	0.048 (0.213)	0.056 (0.231)	0.746	0.132	0.630	0.178	0.165	0.881
Moderately severe depression (15-19)	0.012 (0.107)	0.010 (0.097)	0.021 (0.145)	0.019 (0.137)	0.303	0.193	0.791	0.765	0.122	0.199
Severe depression (20-27)	0.012 (0.107)	0.010 (0.097)	0.008 (0.091)	0.009 (0.094)	0.767	0.942	0.928	0.841	0.884	0.717
Total number of problems felt at least several days (0-9) – female	2.469 (2.879)	2.057 (2.674)	2.211 (2.859)	2.449 (2.974)	0.947	0.213	0.406	0.182	0.590	0.356

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.6.4: Primary male depression, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Severity score - male	2.930 (4.079)	2.646 (4.015)	2.683 (4.041)	2.978 (4.312)	0.918	0.477	0.516	0.528	0.934	0.572
No depression	0.461 (0.499)	0.470 (0.499)	0.471 (0.500)	0.460 (0.499)	0.978	0.844	0.828	0.876	0.985	0.861
Minimal depression	0.285 (0.452)	0.305 (0.461)	0.310 (0.463)	0.273 (0.446)	0.720	0.320	0.285	0.518	0.877	0.454
Mild depression	0.178 (0.383)	0.173 (0.379)	0.152 (0.359)	0.190 (0.392)	0.715	0.614	0.190	0.893	0.489	0.399
Moderate depression	0.054 (0.227)	0.029 (0.169)	0.052 (0.223)	0.048 (0.214)	0.727	0.246	0.815	0.112	0.154	0.912
Moderately severe depression	0.019 (0.135)	0.011 (0.105)	0.004 (0.067)	0.019 (0.135)	0.995	0.352	0.061	0.327	0.206	0.048
Severe depression	0.003 (0.056)	0.011 (0.105)	0.010 (0.102)	0.011 (0.103)	0.153	0.959	0.957	0.237	0.925	0.228
Total number of problems felt at least several days (0-9) - male	2.331 (2.822)	2.141 (2.675)	2.155 (2.732)	2.366 (2.903)	0.915	0.472	0.492	0.558	0.964	0.580

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.6.5: Use of antenatal and postnatal services during the last pregnancy, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Primary female received ANC	0.767 (0.423)	0.748 (0.434)	0.775 (0.418)	0.751 (0.433)	0.691	0.958	0.552	0.616	0.470	0.824
Primary female went to health facility for ANC 4+ times	0.316 (0.465)	0.320 (0.467)	0.301 (0.459)	0.283 (0.451)	0.440	0.375	0.662	0.928	0.628	0.707
Primary female took iron and folic acid supplements	0.421 (0.494)	0.443 (0.497)	0.481 (0.500)	0.404 (0.491)	0.719	0.367	0.075	0.630	0.376	0.192
Primary female received nutrition information/counseling	0.594 (0.491)	0.526 (0.500)	0.564 (0.496)	0.524 (0.500)	0.113	0.963	0.370	0.113	0.382	0.467
Primary female received breastfeeding information	0.556 (0.497)	0.511 (0.500)	0.579 (0.494)	0.516 (0.500)	0.377	0.905	0.140	0.318	0.110	0.599
Birth in a medical facility	0.364 (0.481)	0.275 (0.447)	0.292 (0.455)	0.296 (0.457)	0.151	0.683	0.939	0.072	0.724	0.113
Primary female received vitamin A supplement at or soon after birth	0.213 (0.410)	0.168 (0.374)	0.241 (0.428)	0.185 (0.389)	0.417	0.589	0.106	0.128	0.014	0.386
Primary female received breastfeeding help after giving birth	0.259 (0.438)	0.236 (0.425)	0.303 (0.460)	0.241 (0.428)	0.632	0.893	0.105	0.526	0.070	0.241

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.6.6: Child health history, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Index child received dose of Vitamin A in the past 6 months	0.287 (0.453)	0.293 (0.455)	0.288 (0.453)	0.272 (0.445)	0.655	0.525	0.634	0.855	0.879	0.977
Index child received receive any micronutrient powder in the past 6 months	0.058 (0.234)	0.065 (0.247)	0.069 (0.254)	0.058 (0.233)	0.978	0.643	0.504	0.699	0.816	0.565
Index child's weight was measured in the past 3 months	0.271 (0.445)	0.279 (0.449)	0.271 (0.445)	0.274 (0.446)	0.960	0.899	0.950	0.862	0.851	0.990
Index child's height was measured in the past 3 months	0.243 (0.429)	0.246 (0.431)	0.237 (0.425)	0.262 (0.440)	0.650	0.703	0.549	0.937	0.821	0.886
Index child's MUAC was measured in the past 3 months	0.289 (0.453)	0.280 (0.449)	0.315 (0.465)	0.291 (0.455)	0.956	0.800	0.586	0.829	0.397	0.518
Child feeding information was given at the time of measurement	0.704 (0.458)	0.687 (0.465)	0.691 (0.463)	0.727 (0.446)	0.610	0.458	0.413	0.761	0.936	0.781
Index child identified as severely malnourished in past 6 months	0.087 (0.283)	0.099 (0.298)	0.112 (0.316)	0.077 (0.267)	0.651	0.363	0.107	0.642	0.553	0.264
Received any referral to a facility to receive treatment for severe malnutrition	0.079 (0.272)	0.107 (0.311)	0.136 (0.345)	0.182 (0.389)	0.151	0.317	0.553	0.613	0.642	0.328
Index child had fever in past 2 weeks	0.113 (0.317)	0.125 (0.331)	0.138 (0.345)	0.115 (0.320)	0.950	0.704	0.410	0.617	0.593	0.323
Index child had cough/cold in past 2 weeks	0.150 (0.358)	0.133 (0.340)	0.190 (0.393)	0.118 (0.323)	0.237	0.564	0.013	0.531	0.051	0.193
Index child had fast breathing/shortness of breath in past 2 weeks	0.041 (0.198)	0.031 (0.173)	0.031 (0.174)	0.025 (0.157)	0.254	0.576	0.568	0.491	0.969	0.524
Index child had diarrhea in past 2 weeks	0.090 (0.286)	0.081 (0.274)	0.075 (0.264)	0.092 (0.289)	0.920	0.595	0.362	0.642	0.708	0.382
Index child received ORS when they had diarrhea	0.545 (0.502)	0.409 (0.495)	0.482 (0.504)	0.493 (0.504)	0.523	0.338	0.916	0.128	0.481	0.527
Index child received zinc to treat the diarrhea	0.246 (0.434)	0.095 (0.296)	0.167 (0.377)	0.254 (0.439)	0.916	0.025	0.309	0.015	0.302	0.315

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

Table 6.6.7: Childcare activities, by treatment arm

	Mean and Standard Deviation				P-Value					
	T1:L*+N*	T2:L*+N	T3:L+N*	C:Control	T1 v C	T2 v C	T3 v C	T1 v T2	T2 v T3	T1 v T3
Primary female										
read books or looked at picture books with index child	0.017 (0.131)	0.016 (0.127)	0.020 (0.140)	0.021 (0.145)	0.698	0.571	0.883	0.905	0.634	0.779
told stories index child	0.061 (0.239)	0.051 (0.221)	0.061 (0.239)	0.078 (0.268)	0.403	0.126	0.378	0.581	0.532	0.990
sang songs to or with index child	0.207 (0.405)	0.195 (0.396)	0.206 (0.405)	0.243 (0.429)	0.318	0.191	0.305	0.725	0.736	0.985
took index child outside the home	0.595 (0.491)	0.621 (0.485)	0.583 (0.493)	0.631 (0.483)	0.336	0.789	0.204	0.491	0.318	0.743
played with index child	0.771 (0.421)	0.802 (0.399)	0.780 (0.415)	0.792 (0.406)	0.501	0.728	0.689	0.293	0.438	0.775
named, counted, or drew things with or for index child	0.162 (0.369)	0.160 (0.367)	0.141 (0.348)	0.154 (0.361)	0.828	0.872	0.682	0.945	0.548	0.533
prepared food for index child who is not exclusively breastfed	0.765 (0.424)	0.832 (0.374)	0.764 (0.425)	0.772 (0.419)	0.784	0.024	0.776	0.002	0.006	0.971
physically fed index child who is not exclusively breasted	0.797 (0.403)	0.845 (0.362)	0.800 (0.400)	0.807 (0.395)	0.649	0.106	0.763	0.024	0.049	0.891
gave index child a bath	0.791 (0.407)	0.846 (0.361)	0.818 (0.386)	0.772 (0.420)	0.548	0.020	0.149	0.078	0.385	0.385
cared for the index child when they were sick	0.854 (0.353)	0.838 (0.369)	0.858 (0.350)	0.854 (0.353)	1.000	0.691	0.934	0.698	0.630	0.936
ate a meal together with index child	0.653 (0.476)	0.749 (0.434)	0.698 (0.460)	0.678 (0.468)	0.463	0.011	0.526	0.002	0.071	0.191
Primary male										
read books or looked at picture books with index child	0.030 (0.171)	0.023 (0.151)	0.020 (0.141)	0.030 (0.171)	0.989	0.505	0.349	0.548	0.752	0.396
told stories index child	0.063 (0.243)	0.061 (0.240)	0.051 (0.220)	0.065 (0.247)	0.912	0.819	0.364	0.917	0.495	0.462
sang songs to or with index child	0.117 (0.322)	0.101 (0.301)	0.102 (0.303)	0.130 (0.337)	0.626	0.201	0.220	0.524	0.961	0.554
took index child outside the home	0.398	0.398	0.404	0.407	0.831	0.832	0.933	0.997	0.888	0.888

	(0.490)	(0.490)	(0.491)	(0.492)						
played with index child	0.558	0.619	0.560	0.584	0.525	0.364	0.516	0.158	0.136	0.964
	(0.497)	(0.486)	(0.497)	(0.493)						
named, counted, or drew things with or for index child	0.099	0.102	0.122	0.106	0.795	0.883	0.549	0.904	0.464	0.416
	(0.299)	(0.303)	(0.328)	(0.308)						
prepared food for index child who is not exclusively breastfed	0.078	0.091	0.094	0.083	0.785	0.680	0.540	0.487	0.868	0.353
	(0.268)	(0.287)	(0.292)	(0.275)						
physically fed index child who is not exclusively breasted	0.336	0.349	0.356	0.375	0.318	0.511	0.600	0.767	0.862	0.616
	(0.473)	(0.477)	(0.479)	(0.485)						
gave index child a bath	0.086	0.105	0.117	0.107	0.256	0.906	0.644	0.380	0.611	0.159
	(0.280)	(0.307)	(0.321)	(0.310)						
cared for the index child when they were sick	0.367	0.406	0.449	0.356	0.864	0.438	0.177	0.525	0.514	0.214
	(0.483)	(0.492)	(0.498)	(0.480)						
ate a meal together with index child	0.389	0.447	0.440	0.446	0.140	0.976	0.884	0.150	0.866	0.233
	(0.488)	(0.498)	(0.497)	(0.497)						

Notes: Estimates from the DFSA SPIR Baseline Survey sample. Standard deviations are in parentheses. P-value is calculated from the test of difference of means between each pair of treatment arms. Standard errors are clustered at the kebele level. Missing responses on any question were treated as an incorrect response.

7. Conclusion

The SPIR impact evaluation contributes to the project learning agenda by contributing evidence on the impact of its integrated programming on household livelihoods, food security and nutrition and tests several innovative extensions to the SPIR interventions to learn if the impacts of the project can be further expanded to other outcomes, including access to markets, aspirations, maternal depression and intimate partner violence.

This baseline survey report provides an overview of the SPIR project interventions, summarizes the impact evaluation and sample design and describes the baseline fieldwork. The report then summarizes the main outcome variables in the study as well as many other contextual variables that will contribute to our understanding about where the project works best and why it works. This summary is provided to inform the context and its appropriateness to studying SPIR as a graduation model social protection program. In addition, this report summarizes a series of balancing tests of the difference in means of key outcomes and contextual variables across treatment arms conducted to explore the extent to which the randomized assignment to treatment yielded Treatment groups and a Control group that are similar, so that the Control can serve as an effective counterfactual.

The baseline survey provides several important lessons about the context that will shape the effect of SPIR on livelihoods, nutrition and related outcomes. First, the setting and sample are well suited to study the potential to expand the impact of the PSNP4 cash and food transfers through additional graduation-model programming around improving livelihoods, nutrition, the role of women in intrahousehold decisionmaking and markets and resilience to shocks. All households in the evaluation sample (in all four treatment arms) are PSNP4 beneficiaries, as planned. Also, these households are poor, with 45 percent living below the $1.90 per person per day international poverty line and 22 percent living below the $1.25 extreme poverty line. Regarding child nutrition, 37 percent of children under age 3 years are stunted (HAZ<-2) and 14 percent are wasted (WHZ<-2). The proportion of children under 3 with moderate acute malnutrition is 16 percent and 6 percent suffer from severe acute malnutrition. Diets are quite limited in general and appear to be low in nutrient-rich foods. The period leading up to the baseline survey was characterized by modest food insecurity.

In addition, there are important regional differences in the context that will help inform in which context the SPIR interventions work best. Oromia is noticeably poorer than Amhara, as shown in estimates of poverty prevalence calculated from household consumption. Oromia also has worse food security than Amhara and mostly lower asset holdings. However, child malnutrition is higher in Amhara than in Oromia. This is consistent with recent evidence that the malnutrition burden may not closely overlap the areas with greatest poverty (Brown et al. 2018).

The learning agenda team will use the information provided here to inform the design of the midline survey to be carried out later in 2019. The midline survey will help assess performance of SPIR on delivery against its objectives and will provide preliminary estimates on of impact on outcomes that are likely to change or improve after a short period of intervention.

References

Aizer, Anna. 2011. Poverty, Violence, and Health: The Impact of Domestic Violence During Pregnancy on Newborn Health. *The Journal of Human Resources* 46 (3): 518-538.

Banerjee A, Duflo E, Goldberg N, Karlan D, Osei R, Pariente W, Shapiro J, Thuysbaert B, and C.Udry. 2015. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236): 1260799-1-15.

Bass J, Neugebauer R, Clougherty KF, Verdeli H, Wickramaratne P, Ndogoni L, Speelman L, Weissman M, and P Bolton. Group interpersonal psychotherapy for depression in rural Uganda: 6-month outcomes. *The British Journal of Psychiatry* 188, no. 6 (2006): 567-573.

Berhane G, Gilligan DO, Hirvonnen K, Hoddinott J, Kumar N, Taffesse AS. 2018. *The Productive Safety Nets Programme 4 – Midline Survey 2018: Highland Outcomes Report.* Washington DC: International Food Policy Research Institute.

Berhane G, Gilligan DO, Hirvonnen K, Hoddinott J, Kumar N, Taffesse AS, Abay K, Tadesse F, Yohannes Y, Worku I, Sabates-Wheeler R, Lind J, Birch I, Kohnstamm S, and M Tefera. 2016. *The Productive Safety Nets Programme IV – Baseline Survey Report 2016.* Washington DC: International Food Policy Research Institute.

Berhane G, Gilligan DO, Hoddinott J, Kumar N and AS Taffesse. 2014. Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Program. *Economic Development and Cultural Change* 63(1): 1-26.

Bernard T., Dercon S., Orkin K., Taffesse A. S. October 2017. 'The Future in Mind: Long-Run Impact of an Aspirations Intervention in Rural Ethiopia.' Washington DC: International Food Policy Research Institute.

Black, RE, Victora, CG, SP Walker, and Maternal and Child Nutrition Study Group. 2013. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *The Lancet*. https://doi.org/http://dx.doi.org/10.1016/S0140-6736(13)60937-X.

Black M.M., Baqui A.H., Zaman K., El Arifeen S., Black R.E. 2009. Maternal depressive symptoms and infant growth in rural Bangladesh. *American Journal of Clinical Nutrition* 89(3):951S–957S.

Bolton, P, Bass, J, Betancourt, T, Speelman, L, Onyango, G, Clougherty, KF, Neugebauer, R, Murray, L and H Verdeli. 2007. Interventions for depression symptoms among adolescent survivors of war and displacement in northern Uganda: a randomized controlled trial. *JAMA* 298(5): 519-527.

Bolton, P, Bass, J, Betancourt, T, Speelman, L, Onyango, G, Clougherty, KF, Neugebauer, R, Murray, L and H Verdeli. 2003. Group interpersonal psychotherapy for depression in rural Uganda: A randomized controlled trial." *JAMA* 289(23): 3117-3124.

Brown, C, M Ravallion and D van de Walle. *Forthcoming*. Most of Africa's nutritionally-deprived women and children are not found in poor households. *Review of Economics and Statistics*.

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

Central Statistical Agency (CSA) [Ethiopia] and ICF. 2016. Ethiopia Demographic and Health Survey 2016. Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF.

Devries, K M, J Y T Mak, C García-Moreno, M Petzold, J C Child, G Falder, S Lim, L J Bacchus, R E Engell, L Rosenfeld, C Pallitto, T Vos, N Abrahams, and C H Watts. 2013. "The Global Prevalence of Intimate Partner Violence Against Women." Science 340(6140): 1527–1528.

Ellsberg, Mary, Henrica AFM Jansen, Lori Heise, Charlotte H Watts, and Claudia Garcia-Moreno. 2008. "Intimate Partner Violence and Women's Physical and Mental Health in the WHO Multi-Country Study on Women's Health and Domestic Violence: An Observational Study." The Lancet 371(9619): 1165–72.

Filmer, Deon and Lant Pritchett. 2001. Estimating wealth effets without expenditure data—or tears: An application to educational enrollments in states of India. *Demography* 38(1): 115-132.

Fulu, Emma, Stephanie Miedema, Tim Roselli, Sarah McCook, Ko Ling Chan, Regine Haardörfer, and Rachel Jewkes. 2017. "Pathways between Childhood Trauma, Intimate Partner Violence, and Harsh Parenting: Findings from the UN Multi-Country Study on Men and Violence in Asia and the Pacific." The Lancet. Global Health 5(5): e512–22.

Hasselmann, Maria H, and Michael E Reichenheim. 2006. "Parental Violence and the Occurrence of Severe and Acute Malnutrition in Childhood." Paediatric and Perinatal Epidemiology 20(4): 299–311

Haushofer J and J Shapiro. 2016. The short-term impacts of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya. *Quarterly Journal of Economics* 131(4): 1973-2042.

Hoddinott, J, Taffese, AS, Berhane, G, Gilligan DO, Hirvonen, K, Kumar, K, Lind, J, Sabates-Wheeler, R, Tefera, M and Y Yohannes. 2017. The Ethiopian Productive Safety Net Programme: Lessons learned from 10 years of evaluation. Manuscript. Washington DC: International Food Policy Research Institute.

Hoddinott, John, Jere R Behrman, John A Maluccio, Paul Melgar, Agnes R Quisumbing, Manuel Ramirez-Zea, Aryeh D Stein, Kathryn M Yount, and Reynaldo Martorell. 2013. "Adult Consequences of Growth Failure in Early Childhood." *The American Journal of Clinical Nutrition* 98 (5): 1170–1178.

Kapiga, Saidi, Sheila Harvey, Abul Khalie Muhammad, Heidi Stockl, Gerry Mshana, Ramadhan Hashim, Christian Hansen, Shelley Lees, and Charlotte Watts. 2017. "Prevalence of Intimate Partner Violence and Abuse and Associated Factors among Women Enrolled into a Cluser Randomised Trial in Northwestern Tanzania." *BMC Public Health* 17(190): 1–11.

Karamagi, Charles AS, James K Tumwine, Thorkild Tylleskar, and Kristian Heggenhougen. 2007. "Intimate Partner Violence and Infant Morbidity: Evidence of an Association from a Population-Based Study in Eastern Uganda in 2003." *BMC Pediatrics* 7(1): 34.

Koenen, Karestan C, Terrie E Moffitt, Avshalom Caspi, Alan Taylor, and Shaun Purcell. 2003. "Domestic Violence Is Associated with Environmental Suppression of IQ in Young Children." *Development and Psychopathology* 15(2): 297–311.

McIntosh, Craig and Andrew Zeitlin. 2018. Benchmarking a child nutrition program against cash: Experimental evidence from Rwanda. IPA Working Paper.

McKenzie, D. 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics* 99(2): 210-221.

Pollak, Robert A. 2004. "An Intergenerational Model of Domestic Violence." *Journal of Population Economics* 17(2): 311–29

Taffesse, Alemayehu Seyoum, and Fanaye Tadesse. 2017. Pathways Less Explored—Locus of Control and Technology Adoption," *Journal of African Economies* 1–37.

Victora, Cesar Gomes, Mercedes de Onis, Pedro Curi Hallal, Monika Blössner, and Roger Shrimpton. 2010. "Worldwide Timing of Growth Faltering: Revisiting Implications for Interventions." *Pediatrics* 125 (3): e473–80.

Wachs TD, Black MM, Engle PL. Maternal depression: a global threat to children's health, development, and behavior and to human rights. *Child Development Perspectives* 2009; 3: 51-9.

WHO. 2016.Ethical and safety recommendations for intervention research on violence against women. Building on lessons from the WHO publication *Putting women first: ethical and safety recommendations for research on domestic violence against women.* Geneva:World Health Organization.

———. 2000. *The Management of Nutrition in Major Emergencies*. Geneva: World Health Organization.

———. 2006. "WHO Child Growth Standards."

———. 2008. "Indicators for Assessing Infant and Young Child Feeding Practices."

Yount, Kathryn M, Ann M DiGirolama, and Usha Ramakrishnan. 2011. "Impacts of Domestic Violence on Child Growth and Nutrition: A Conceptual Review of the Pathways of Influence." *Social Science & Medicine* 72(9): 1534–54.

Appendix A: SPIR Baseline Survey Questionnaire

[The baseline survey questionnaire is not included in this version of the baseline report in order to keep the file size manageable. The complete version of this report, including the questionnaire, is available upon request.]