The Productive Safety Net Programme IV

End-line Outcomes Report (2021) – Lowlands

[a] International Food Policy Research Institute

[b] Cornell University

February 6, 2022

Table of Content

List of Tables

List of Figures

Executive Summary

This report documents the impact of the Productive Safety Nets Programme (PSNP) in two Lowland regions of Ethiopia – Afar and Somali. It updates material presented in previous assessments of program impact and addresses the outcome-related questions listed in the terms of reference for this study. A summary of the answers to these research questions is provided below.

RQ17: How are the transfers used?

- The results indicate that nearly all households in lowland areas spend cash transfers to buy food and about 50 percent use them to buy nonfood items.
- Using the transfers to help other households is reported by 13 percent of households.
- Households allocated over three-quarters of their cash transfers to buy food and about 17 percent to buy nonfood items. The share of cash transfers given to other households is 5 percent.
- Over three-quarters (78 percent) of households report storing their food transfers and a little less than 30 percent report selling it.
- On average, households stored 67 percent of their food transfer, sold 15 percent, and gave 11 percent to other households.

RQ1: To what extent has the PSNP improved food security (including dietary diversity) among households participating in the program?

- Substantial heterogeneity arises in the mean food gap between Afar and Somali. Between baseline and end-line, the mean food gap decreased for all households in Afar and for PSNP households in Somali, but increased for non-PSNP households in Somali. Specifically, among PSNP beneficiaries, the mean food gap declined by 0.4 months in Afar and 0.6 months in Somali. The mean food gap also declined by 0.2 months among non-PSNP households in Afar, but increased by 1.2 months (an increase of 63 percent) among non-PSNP households in Somali.
- The overall mean food gap is also higher in Somali than in Afar, regardless of PSNP status.
- Between baseline and end-line, the share of households reporting a food gap of three months and above declined by 2 percentage points for PSNP households, but increased by 6 percentage points for non-PSNP households.
- Female-headed non-PSNP households in Afar and Somali tend to report a higher food gap than female- and male-headed PSNP households in their respective regions.
- For both PSNP and non-PSNP households in Afar and Somali, the mean household dietary diversity score (HDDS) declined consistently between baseline and end-line.
- PSNP households in Somali consume a slightly more diverse diet (as measured by the number of food groups) than those in Afar. Mean HDDS declined for PSNP households in both regions.

- Close to 90 percent of households consume cereals, while the proportion of households consuming animal-sourced foods and fruits is low. No noticeable difference is observed between PSNP and non-PSNP households in the food group types from which they consume.
- Finally, impact estimates show that participation in the PSNP did not have a statistically significant impact on the food gap or HDDSs. It is important to read these results in the context of several unfavorable macroeconomic conditions observed in the last three years (including COVID-19, instability, droughts, floods, and pests), some of which had broader implications for program performance (for example, poor payment performance) as outlined in the Performance Report chapters and discussion of household-level shocks (for example, drought, flood, and pests).

RQ2: Has PSNP impacted household resilience to shocks?

- Mean Tropical Livestock Units (TLU) held declined between baseline and end-line for all households, but more dramatically so for non-PSNP households – a decline of 43 percent for non-PSNP but only 9 percent for PSNP households. Regional disaggregation also shows a dramatic decline in mean TLU for non-PSNP households in Afar (by 42 percent) and Somali (by 48 percent), but for PSNP households in both regions, the decline was only 8 percent. Mean value of livestock assets followed the same trends.
- Overall, the mean values of productive assets for both rounds and for all household types are lower in Afar than in Somali.
- At end-line, a household from Afar owned a higher total value of assets (28,000 Birr on average) than a corresponding household from the Somali sample (18,000 Birr on average).
- Female-headed households owned consistently lower mean TLU than male-headed households. Overall, at baseline female-headed households owned 45 percent lower mean TLU than male-headed PSNP households. This gap widened further in 2021 when, on average, female-headed PSNP households owned 60 percent less TLU than male-headed households.
- Regardless of their PSNP status, female-headed households in the Lowlands owned a lower mean value of productive assets than their male-headed counterparts.
- Drought or lack of moisture stands out as the most important shock for all households in the Lowlands, regardless of their PSNP status. Between 60–70 percent of all households reported drought as most important.
- Little difference arises between PSNP and non-PSNP households with regard to the proportion of households reporting distress sales. Among the three asset types, distress sales of livestock for food and emergency cash needs stands out as the highest reported (for about 8–12 percent of all households).
- Impact estimates show that the PSNP had no statistically significant impact on any of the household assets discussed in this chapter, mainly livestock TLU, real value of livestock, real value of durable assets, and real value of total assets.

RQ3: Has PSNP reduced poverty?

- PSNP households' perception of their economic standing is lower than that of non-PSNP households at baseline and at end-line. We observe some movement in this perception over time. The percentage reporting "little poorer than most" and "the poorest" rose in both groups. Among non-PSNP households, the percentage self-reporting as "the poorest" more than doubled during this time.
- In terms of absolute poverty, we observe that at baseline, 46 percent of PSNP households considered themselves as destitute or poor. The corresponding percentage among non-PSNP households was slightly lower, at 43 percent. These proportions remained almost the same over the time covered. Nevertheless, it is noteworthy that among PSNP households, the group of destitute households increased from 9 percent to 13 percent, while it climbed from 5 percent to 10 percent among non-PSNP households.
- Male-headed households are better off than female-headed households from the perspective of subjective poverty. But over time, both groups saw increases in the fraction of households identifying themselves as destitute. On a positive note, the percentage of households claiming to be comfortable rose for both groups, albeit from a very low base.
- The two regions had similar levels of subjective poverty at baseline. The end-line is associated with a higher proportion of households reporting to be destitute in both regions, while the share of those identifying themselves as comfortable grew in the Somali region.
- About one-half of households (irrespective of PSNP status) were able to obtain 100 Birr at baseline. Over time, we observe small changes in this ability among non-PSNP households. But among PSNP households, the ability to obtain emergency funds increased appreciably.
- Comparing male- and female-headed households, we find that a larger percentage of male-headed households were able to obtain emergency funds at baseline, but by end-line both groups experienced improvements in their ability to raise emergency funds and the gaps reduced significantly.
- Little regional variation existed in the ability to raise emergency funds at baseline. By end-line, households in Afar showed an improvement in their ability to raise emergency funds, but we do not see such improvement in Somali.
- Compared to two years ago, 16 percent of households overall reported doing much worse at end-line, 18 percent reported doing a little worse, 44 percent reported doing the same, and 18 percent reported doing a little better. We do not observe significant variations across different groups.
- We do not find any impact of the PSNP on subjective poverty.

Descriptive results on agency

- PSNP and non-PSNP households are very similar in their perception of having full control over their life. Over time, we observe some deterioration in both groups.

- In terms of speaking up in public – on matters relating to infrastructure, payments in the public works, if some aspects of the PSNP are not implemented fairly or correctly – more than 40 percent of households are not at all comfortable, about 20–25 percent are comfortable, and the remaining 34–36 percent would speak up with difficulty.

RQ11: Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?

- The report presented descriptive results on labor allocation and work intensity in the seven days prior to the end-line survey. The activities included are agriculture (including livestock- and fishing-related activities), nonagricultural business, casual/part-time labor, wage or salaried labor, and work on PSNP Public Works.
- Almost one-half (46 percent) of households reported not having spent any time on any activities related to agricultural or nonagricultural work, casual labor, salaried work, or Public Works in the seven days prior to the survey. We observe differences by region, sex of household head, and PNSP status.
- PSNP households are less likely than non-PSNP households to report not having worked in the past seven days.
- On average, households were engaged in less than one out of five activities in the last seven days.
- On average, 43 percent of households reported being engaged in agricultural activities in the last seven days. The rest of the activities are reported by a small percentage of households.
- PSNP Public Works are reported by 9 percent of households; as expected, this is concentrated among PSNP households.
- PSNP households spent a larger number of hours on various activities in the last seven days compared to non-PSNP households.
- Given the data constraints, we cannot answer the research question, **RQ11: Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?** But we do find that PSNP households are more likely to be working and work more hours compared to non-PSNP households. This may reflect their relative economic status, which may require PSNP households to work more. However, it may also suggest that the PSNP is not leading to a reduction in work effort.

Chapter 1: Introduction

1.1 Background

Beginning in 2005, the Government of Ethiopia and a consortium of donors implemented a new response to chronic food insecurity in rural Ethiopia. Rather than annual appeals for assistance and ad hoc distributions, the Productive Safety Nets Programme (PSNP) was established.

The PSNP "…provides transfers to the food insecure population … in a way that prevents asset depletion at the household level and creates assets at the community level". Unlike annual emergency appeals, it was conceived as a multi-year program to provide recipients with predictable and reliable transfers. The PSNP uses a mix of geographic and community-based targeting to identify beneficiaries. Approximately 80 percent of participants receive six months of employment on labor-intensive public works projects. These emphasize reversing environmental degradation, improving water control and improving road access. The remainder, largely households whose primary income earners are elderly or disabled, receive unconditional transfers. Payments are made in both food and cash.

Despite its achievements, the precarious nature of livelihoods in localities means that a social protection intervention like the PSNP is still required in these areas. Additional efforts are needed to integrate the program with nutrition and agricultural extension services. Clear awareness of these on the part of the Ethiopian government and its development partners led to the design and adoption of PSNP 4.

PSNP 4's overall Project Development Objective is to achieve 'Increased access to safety net and disaster risk management systems, complementary livelihoods services and nutrition support for food insecure households in rural Ethiopia' (World Bank 2014: 21). It will attempt to achieve this through: 1. Support for building core instruments and tools of social protection and DRM systems; 2. Delivery of safety net and enhanced access to livelihoods services for vulnerable rural households; 3. Improved program management and institutional coordination. A prominent theme of PSNP 4 is the desire to integrate frameworks and move towards a 'systems approach' (GFDRE 2014). This requires the necessary administrative structures and institutional capabilities to consolidate overlapping areas into a more integrated and predictable safety net.

A number of innovations in the design of the PSNP 4 contribute to these objectives. First, program support will be organized around the idea of 'livelihood pathways', with packages of support (transfers, technical assistance, access to credit, training) tailored for different categories of chronically food insecure households. The incorporation of a livelihoods component in PSNP 4 builds on the architecture and delivery mechanisms established under the Household Asset Building Programme (HABP), which no longer exists as a separate program. The provision of new livelihood transfers (start-up capital), drawing on practice from other contexts, it is hoped will help to increase the prospects of graduation, which have been disappointing thus far. It is intended by this merging that support for building livelihoods and supporting graduation is closely integrated with other program components that work well, including the delivery of transfers and public works implementation. A second innovation is the goal of shifting the Permanent Direct Support caseload to

the Ministry of Labor and Social Affairs (MOLSA). This recognizes that there are categories of households that have very little or no prospect of 'graduation' and will always need some level of assistance, such as the old, chronically sick, or disabled. One of the strengths of the PSNP was establishing robust delivery systems and capacity at all administrative levels down to the kebele-level. It is hoped in the PSNP 4 that the operational capacity of the MOLSA will be expanded at the woreda and kebele-levels. A third innovation is expansion in the program's coverage to cover chronically food insecure households in 92 more woredas. This is a significant expansion in the program's reach; thus, it will be important to assess the PSNP's performance in new woredas where the program is just being introduced.

1.2 Objectives, research questions, and focus

1.2.1 Objectives of the evaluation

This impact evaluation uses a mix of quantitative and qualitative research methods to meet the three objectives of the impact evaluation, summarized as:

- assess progress in program performance across all components and implementing agencies;
- provide a rigorous assessment of the impact of the PSNP on wellbeing and livelihoods of households, with regards to (a) provision of safety net transfers in cash and food, (b) provision of livelihood technical support and transfers, and (c) promotion of linkages to nutrition and health programs; and
- provide insights into why and how these impacts were achieved.

This report delivers on the second objective by providing a rigorous assessment of the impact of the PSNP on poverty, resilience to shocks, food security and human development indicators

1.2.2 Research questions and focus

This report presents analysis that addresses the following research questions:

Table 1.1: Research questions addressed in this report

#	Research question
RQ1	To what extent has the PSNP improved food security (including dietary diversity) among households participating in the program?
RQ2	Has PSNP impacted household resilience to shocks?
RQ3	Has PSNP reduced poverty?
RQ11	Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?
RQ12	How does the shock responsive component of PSNP (the federal contingency budget) protect people against covariate shocks?
RQ17	Use of transfers

1.3 COVID-19 and the PSNP

The COVID-19 pandemic that emerged in late 2019 has affected the lives and livelihoods of millions of households globally. Policies put in place to battle the pandemic have created further economic

hardships, possibly leading to the worst global economic crisis in decades (World Bank, 2020). Stay-at-home measures and social distancing restrictions put in place to curb the spread of the virus have led to dramatic declines in economic activity and disruptions of social life. Developing countries are likely to be the worst affected in this crisis. Ethiopia was no exception. As of June 4, 2021, about 272 thousand cases and 4,185 deaths were reported in Ethiopia. Soon after the first COVID-19 case was reported in Ethiopia on March 13, 2020, the pandemic started to take a toll on economic life. A state of emergency declared on April 8th to stem the spread of the virus expanded a set of related restrictions to include bans on cross-border movements except 'essential' cargo transport, restrictions on public gatherings and movement of people, imposition of partial stay at home orders for workers, and the closing of schools. Some regional states have also put in place stricter restrictions on the movement of people and goods that may potentially disrupt the functioning of markets and affect economic lives. Those measures to prevent the spread of the virus, might have caused disruptions in the livelihoods of poor households and those reliant on farming in rural Ethiopia. Although the more stringent measures were quickly removed in the case of Ethiopia and people remained cautiously optimistic, public movements and gatherings of more than a certain level remained restricted.

Ethiopia is one of few Sub-Saharan African countries with a large safety net programme, PSNP, in place with the potential to tackle the negative effects of COVID-19. We undertook a phone survey in 2021 to examine the implications of the pandemic on households in rural Ethiopia and how these differ by PSNP status. The detailed findings are presented in Berhane et al 2021, here we refer to some key findings as they pertain to the findings of the outcomes report.

Overall, movement wise, non-PSNP households are more likely to report presence of mobility restrictions and lockdown measures. However, while about two-third of all households reported leaving their house in the last 7 days, non-PSNP respondents were slightly more likely to have done so as compared to PSNP respondents. With regards to access to health services, a majority reported they were being able to go to hospital/medical facility whenever they needed to – however, non-PSNP households are more likely to report this than PSNP households. When asked about the most disruptive events since the start of the pandemic - over 60% reported higher food prices, 43% reported unemployment or loss of income, 39% said shortages in food supply, about 30% mentioned school closures, and 28% were affected by travel restrictions. PSNP households were more likely to report loss of incomes as compared to non-PSNP households.

Among those that received public works, a vast majority reported that they were requested to carry out public works to receive these transfers after Megabit 2012 (during COVID-19 period). Clearly, this not in line with the COVID-19 protocol that outlined households would not be required to do public works to receive payments. In terms of nature of payments, households have reported receiving transfers as food, cash, and combination of both. There was significant variability in terms of the frequency and size of payments made. Reassuringly, a large majority of respondents reported that the value of transfers received had not changed or even increased from pre-pandemic levels. Only 13% reported decrease in the value of transfers received.

Overall, a large proportion of the households reported earning much less (37%), somewhat less (39%) or about the same (14%). PSNP households are more likely than non-PSNP households to report

receiving somewhat less income. When asked about the ability to sell items in the market compared to pre-COVID, 55% of households report it being harder, with the PNSP households being more likely than non-PSNP households to report this. Following from a loss income, the 73% of the households reported feeling worried about not having enough food to eat. PSNP households are more likely to report this as compared to the non-PSNP households. About 13% of the households also reported that they went without food for a whole day, this group was also overrepresented by PSNP households.

As a primary coping strategy, 58% of PSNP households and 50% of non-PSNP households have reported consuming poorer quality food in the 30 days prior to the survey. Borrowing money to buy food (63% PSNP vs 36% non-PSNP) and selling productive assets (63% PSNP vs 31% non-PSNP) stand out as the next two most important coping mechanisms followed. Others responses include reducing essential non-food expenditures (43% PSNP vs 38% non-PSNP); lowering health expenditures (34% PSNP vs 31% non-PSNP) and drawing down savings (42% of both PSNP and non-PSNP). Food insecurity worsened despite these coping responses. The likelihood of being food insecure increased respectively by 19.9 and 28.5 percentage points for PSNP and non-PSNP households, while the food gap rose on average by 0.9 months for both groups.

Two important findings are drawn from the findings of the PSNP phone survey report regarding the protective role of the PSNP on food security during the COVID-19 crisis. First, compared to the year before COVID-19, food security has indeed deteriorated during the COVID-19 year: it is estimated that overall the likelihood of becoming food insecure has increased by 37 percentage points and food gap has increased by 1 month. Participation in the PSNP protects households from becoming more food insecure and it is associated with a reduction in the food gap.

Chapter 2: Data and Methods

2.1 Introduction

In this chapter, we describe the quantitative end-line survey, summarize the baseline and end-line data used in the evaluation, and describe the matching methodology used to estimate the impact of the PSNP4 in the Lowlands. The approaches to data collection and evaluation methods used in this study are similar to those used in our previous impact evaluations of the PSNP in the Lowlands. This evaluation is based on quantitative and qualitative data collected in three rounds: a baseline survey in 2016, a midline survey in 2018 and an end-line survey in 2021. Detailed baseline and midline reports have summarized the results of those surveys and the midline report provided intermediate estimates of the impact of the program after less than two years of operation. This end-line report provides the main impact analysis for PSNP4 in the Lowlands, after the program was in operation for four years.

The quantitative end-line survey was originally scheduled for 2021 but was postponed because the COVID-19 pandemic led to a pause in all household survey data collection in Ethiopia, for safety reasons. In 2021, the end-line survey was collected between April and May. This is close to the timing of the 2018 survey which was June/July 2018 (rainy, lean season) but later than the baseline data which were collected in January/February 2016 (dry, postharvest season). The timing of the 2021 surveys was dictated by the need to ensure that PSNP activities had been undertaken in 2021 (these data are needed for many of the research questions addressed in the Performance Report) but also by the need to ensure that data collection was completed well in advance of national elections scheduled for June 2021. The main quantitative end-line survey in the Lowlands was conducted through in-person interviews in Afar and Somali Regions. Data collection took place in April and May, 2021. Data collection for the end-line survey was conducted by the Central Statistics Authority (CSA) with support from IFPRI.

The survey instruments were developed in consultation with the Food Security Coordination Directorate (FSCD) and representatives from the PSNP Donor Coordinating Team (DCT). The final version of the end-line survey instruments were reviewed and approved by these stakeholders.

The design of the quantitative sample was based on careful power calculations conducted to determine the minimum number of sample enumeration areas and households needed to be able to identify impacts of the PSNP4. This involved carefully stratifying the sample between PW and DS households as well as the inclusion of non-beneficiary households into the sample. The sampling strategy (including the statistical power calculations) are described in the inception report and the baseline report and summarized in Appendix A of this end-line report.

The evaluation uses a repeated cross-section design in the Lowlands after it was determined, through consultation with CSA, that tracking households from the baseline sample would be very challenging in Afar and Somali.[1] The total number of households interviewed during the 2016 baseline survey was

[1] Households in the lowlands move in search of water and grazing for their animals. This is particularly true of the PSN$ survey period (February-May). Even when GPS location has been recorded for a household in one

1,983 across 24 woredas and 72 enumeration areas in the Lowlands. A target sample of similar size was targeted for the 2021 end-line survey, based on a new cross-sectional sample of households from the same woredas.

2.2 Sampling

The sampling strategy is carefully described in the inception and baseline reports, and summarized in Appendix A. The passages below offer a brief overview.

Three steps were involved in the selection of households for the PSNP4 baseline. First, the 24 woredas were randomly selected from among the pool of PSNP4 woredas using proportions derived from population size and project coverage. At the second stage, 3 enumeration areas (EAs) were randomly chosen from among EAs in each woreda. The final step was the selection of 28 households from within each EA. This was done based on a fresh listing of households residing within each EA during the baseline in 2016. The listing form used for this purpose gathered information on household current and past PSNP beneficiary status; age and gender of the household head; household land and livestock holdings; and household wealth self-ranking relative to other village residents. Households were then randomly selected from this list until the desired number and composition of households were obtained. To maximize the chance of obtaining a control sample that is as similar as possible to the treatment sample, the non-beneficiary (control) households were chosen from the same EAs and from the bottom four rungs of the subjective wealth ranking.

In 2016, the total number of households interviewed during the baseline survey was 1,983 in Afar and Somali. These households were sampled from 72 enumeration areas (located in 72 kebeles) in 24 woredas. In 2018, a separate cross-section of 1,945 households was interviewed in the same woredas, kebeles and EAs in the Lowlands. The 2021 survey round in the Lowlands reached a new cross-section of 2,084 households from the same 24 woredas and from the same kebeles and EAs.

2.3 End-line survey

Data collection followed a "cascading" design with surveys conducted at the woreda, kebele, and household level. We describe each survey here.

2.3.1 The woreda quantitative survey

The primary purpose of the woreda quantitative survey is to collect information on how the flow of funds from regions to beneficiaries works in practice and the resources used to support those flows. It included the following modules

A. Basic woreda characteristics and infrastructure
B. Staff directly engaged with the Productive Safety Net Program
C. PSNP4, General (including beneficiaries, payment modalities and other transfers or services)
D. Infrastructure and staff specific to the Productive Safety Net Program
E. Contingency budgets
F. Targeting

round, it is not uncommon to find another household in a subsequent round in that location. Considerable similarity in household heads' names add to the difficulty.

G. Cash payments (including payment schedules, attendance sheets, obtaining funds, making payments)
H. Food payments (including payment schedules, attendance sheets, obtaining food, making payments)
I. Comparative experiences with food and cash transfers
J. Humanitarian Food Assistance
K. COVID-19

Interviews were completed by experienced survey supervisors who were instructed to meet with staff associated with the woreda Food Security Office (WFSO) as well as those knowledgeable of the payment system. Ideally, they were supposed to speak with the Head of the Food Security Office, the WOFED chief accountant, the PSNP accountant, and PSNP cashiers. In addition, they were encouraged to seek out and interview individuals knowledgeable about the HABP such as the woreda extension desk leader or the head of the woreda Cooperative Promotion Office.

2.3.2 The quantitative community and price questionnaire

In this questionnaire, the community is defined as the kebele or peasant association (PA). Enumerators were instructed to interview at least five people, perhaps together, who are knowledgeable about the community (e.g., community leaders, PA chairpersons, elders, priests, teachers). They must include at least one member of the Kebele Food Security Task Force and at least one woman and they are told that they may need to meet with other members of the Kebele Food Security Task Force in order to complete some sections of this questionnaire.

As it did in previous years, the community questionnaire covered the following topics: location; physical access and basic services; health services; support for livelihoods, agricultural production and marketing services; wages; prices of food grains in the last year; operational aspects of the PSNP, including questions about the operations of the FSTFs; public works and direct support; the Kebele Appeals Committees; other forms of assistance and Responses to COVID-19. Questions were also asked about the Livelihood Component, on attitudes toward targeting and on moving pregnant women to Temporary Direct Support.

2.3.3 The quantitative household survey and questionnaire

As noted above, the 2016 baseline survey included 1,983 households in Afar and Somali. These households were sampled from 72 kebeles or enumeration areas in 24 woredas. Also, we note that PSNP was not operational in 2 sampled woredas (both in Somali region) in 2016. These two woredas are omitted from the sample for analyses that follow. This reduced the effective size of the baseline sample to 1,803 households in 22 kebeles. The 2021 end-line survey included 2,084 households from the same 22 woredas. Table 2.1 shows the number of households by round and region.

Table 2.1: Number of households interviewed, by round and region

	Lowland	Afar	Somali
2016	1,983	1,028	955
2021	2,084	1,080	1,004

Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

The household questionnaire is a multi-topic instrument collecting information on household participation in the PSNP, their knowledge of PSNP operations and data needed to construct outcome indicators that are required for the Outcomes Report.

Table 2.2 lists the modules and briefly describes the contents of the household questionnaire fielded in the Lowlands.

Table 2.2: Contents of the household questionnaire fielded in the Lowlands

Module	Title	Description
Adult male questionnaire		
M1	Basic household characteristics	This module covers household demographics, current household members; characteristics of the household and the household head; time use; employment; and former household members.
M2	Land, crop and forestry production and disposition	The module captures crop production outcomes over the last 12 months as well as crop production activities relevant to the PSNP livelihoods component.
M3	Household assets, livestock and livestock production	This module collects data on the assets owned by the households. Greater attention is given to livestock and livestock products including ownership, production and sales, and extension since these are likely to be affected by the new PSNP livelihoods component. Information on distressed asset sales is also included.
M4	Income apart from own-agricultural activities and credit and savings	Activities/topics covered in this module include wage employment; own business activities; private transfers; credit for productive purposes; credit for consumption purposes; and savings and access to savings institutions. Also, the spouse (adult female) questionnaire for the highlands is designed to capture the gender differences across these domains/activities.
M5	Access to the PSNP	The module covers access to the PSNP (public works and direct support) during the past year; understanding of PSNP4 operations including, targeting and appeals process, selection of public works projects; other public transfers; the livelihood components; and graduation.
M6	Consumption	This module collects data on households' non-food expenditures and their views on food consumption, including on food security status. Note that the detailed module on food consumption is now part of the spouse (adult female) questionnaire in the highlands. Both head and spouse (female) now answer questions regarding food security.
M7	Health, illness, shocks, poverty perceptions, and decision-making and voice	This module collects data on households' health status; experience of shocks (including COVID-19); their perceptions on poverty; and decision-making and voice.
M8	Nutritional status	Mid-upper arm circumference for all women under age 40; mid-upper arm circumference for all children age 6-59 months.

2.4 Impact evaluation strategy

2.4.1 Overview

The central challenge of any impact evaluation is to estimate impact by comparing outcomes for beneficiaries to the *counterfactual* – what those outcomes would have been had the beneficiaries not received the program. In a randomized controlled trial (RCT) design, the counterfactual is constructed by randomly assigning treatment and control group status between similarly eligible communities or households. When treatment assignment is random, households assigned to the control group are identical (in expectation) to households in the treatment group at baseline, so these control households provide a strong counterfactual. Impacts of the program can be measured as differences in outcomes (or differences in changes in outcomes over time) between the randomly assigned treatment and control households. When it is not possible to implement an RCT or other experimental design, an *identification strategy* must be developed in which the counterfactual is constructed using statistical techniques to create a comparison group of households who are observationally similar to the beneficiary group.

The primary approach in this evaluation *combines difference-in-difference* and *matching methods* for impact evaluation. Matching approaches can be used for programs like the PSNP in which targeting of beneficiaries at the community and household level is conducted by the program and cannot be subjected to randomization and where a regression discontinuity design is infeasible (the use of community targeting means that there is no unique cut-off separating beneficiaries from non-beneficiaries). Members of this research team have used matching methods successfully to evaluate the impact of the PSNP in previous rounds (Berhane et al. 2014, Gilligan, Hoddinott, and Taffesse 2009).

A common approach to estimate program impacts in non-experimental studies is to construct the counterfactual by matching program beneficiaries to non-beneficiaries using observed characteristics measured in the household survey by matching techniques such as propensity score matching (Rosenbaum and Rubin 1983, Smith and Todd 2005) or covariate matching (Abadie and Imbens 2006). Such approaches estimate program impacts as a weighted average of differences in outcomes between beneficiaries and non-beneficiaries in which the weights are constructed using a measure of the degree of similarity of characteristics of households in these two groups. These approaches differ in part by the way that they construct the weights. Heckman, Ichimura and Todd (1997, 1998) show that such matching approaches work well under certain conditions (including that households live in the same areas/markets and that observable characteristics are measured in the same way). These conditions are likely to be met in the PSNP 4 evaluation samples because non-beneficiary households are sampled from PNSP kebeles and woredas.

As described in the inception report, the sampling strategy used in the Highlands (Amhara, Oromia, SNNP and Tigray) and the so-called Lowlands (Afar and Somali) are different. In the Highlands, we have a panel (longitudinal) survey; each household interviewed at the baseline is intended to be re-interviewed in the midline and in the end-line. In the Lowlands, in contrast, we resort to a repeated

cross-sectional design; while the survey localities will remain the same, different households are interviewed in each survey round.

Given these differences in the sampling design, we need to apply somewhat different statistical methods to evaluate impact. In the Highlands, we can construct the counter-factual by matching beneficiary and non-beneficiary households using household and community characteristicsmeasured at the baseline in 2016. Further, the primary outcome variable will be *change* in food gap[2] between baseline and midline/end-line. In the Lowlands, we cannot apply the same strategy becausethe households will not remain the same between the baseline and follow-up. We therefore, opt for a slightly different matching strategy that combines difference-in-difference with matching in a context of repeated cross-section data. More specifically, following Blundell and Dias (2009) and Blundell et al. (2004), we match the treated households at end-line separately to three comparison groups: treated group at the baseline (i.e. before treatment), control group at the baseline, and control group at follow-up. This approach creates the propensity score weights needed to construct the difference in outcomes across these groups. A single propensity score model is estimated on a sample including all four of these groups in which the participation variable is defined as 1 for all treated end-line observations and 0 for all treated baseline, comparison end-line and comparison baseline observations.

Calculating the treatment effect using the repeated cross section Propensity Score Matching (PSM) approach requires doing the following: for each treated end-line outcome observation, subtract the weighted average treated baseline outcome and the difference in the weighted average outcome between end-line and baseline in the comparison group. The average of these differences in weighted averages across treatment and comparison observations is the estimated impact of the program. The repeated cross-section approach to matching estimates each of these four components of the impact estimate separately for each treated observation, calculates the impact estimate and then averages those estimates across all treated observations. The panel approach to matching calculates the impact estimate as the average difference in the outcome for each treated observation from a weighted average of the outcome in the comparison group. As with other PSM models, standard errors for measuring significance levels must be constructed using bootstrapping. For estimated impacts on child nutritional status, child feeding practices, health service utilization and maternal nutrition knowledge, we applied the repeated cross section approach to estimate impacts from the Lowlands sample for two sets of models, one for each comparison group: thus, comparing T vs C1 and comparing T vs C2.

In our application, this strategy translates into conducting the matching three times:

1. PSNP households in 2021 (treated) to non-PSNP households in 2021 (control)
2. PSNP households in 2021 (treated) to PSNP households in 2016 (control)
3. PSNP households in 2021 (treated) to non-PSNP households in 2016 (control)

[2] The food gap is defined as the number of months the household cannot satisfy its food needs. This is based on household's own assessment.

2.4.2 Definition of treatment status

An important feature of the evaluation is the definition of treatment status. PSNP4 has been operating for more than five years. Over that time, it is expected that there would be some variation in the number of years that some households would have participated in the program. There may also be some variation in whether people who believe they are beneficiaries of PSNP4 are actually receiving payments from the program, due to some discrepancy in their beneficiary status or a delay in delivery of payments.

The household questionnaire had a series of questions that asked households whether they had been selected into the PSNP during the survey year or in the previous years. In addition, the households were asked a series of follow-up questions that help us to correctly determine households' beneficiary status, including questions whether they have been selected into public works, direct support; the payments they received and the amount of public works they have been carrying out in the past 18 months.

For the purposes of this analysis, we need to define PSNP beneficiary status in both the baseline and end-line surveys because we collect data on different households in each round. For this analysis, in 2016 (baseline), we define households as PSNP beneficiary households if they reported to have participated in either the public works or the direct support component of the program in 2015/16. Ideally, we would have triangulated this information against data on payments. However, the baseline survey was administered in January-February so only few households had received PSNP payments.

In 2021 (end-line), we define households as PSNP beneficiaries if they reported to have participated in either the public works or the direct support component of the program in 2021 or in 2021. This is the same definition used for the Highlands analysis, so we retained it for the Lowlands as well. We extensively examined whether using payment data would help to identify beneficiaries and we found that adding a criterion based on payments did not substantially alter the beneficiary definition.

2.4.3 Covariate selection

As required for matching, we selected variables for the matching models that are likely to affect both the selection into PSNP (as measured by PSNP status in 2021) and outcomes. This selection is guided by theory, knowledge of how the program functions and our previous impact evaluations of the PSNP. It should be noted that all these variables are measured at baseline – before the fourth phase of PSNP began. Broadly, the household level variables can be categorized into head's characteristics (age, age squared, education level), household demographics (size and dependency ratio), household assets (housing, land, livestock, etc.), housing characteristics, exposure to shocks and community variables. The community level characteristics include number of development agents in the community, road access and community level shocks, and whether the community received humanitarian aid in the past 12 months.

Table 2.3 provides summary statistics for selected basic household characteristics in our sample: household head's characteristics and household size. These are reported by region and for the full

Lowland sample. We note that household heads are 3.4 years older on average in Somali than in Afar and have higher average education by 0.402 years. On average, 17 percent of household heads have any education in Somali and only 12 percent in Afar. Households are also larger in Somali by roughly 0.4 members, due to having slightly larger number of children and more elderly members.

Table 2.3: Summary statistics of household characteristics, by region

	Afar			Somali			All		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Head age, years	1080	40.347	13.018	884	43.736	15.230	1964	41.873	14.154
Head education, years	1058	0.822	2.488	851	1.224	3.051	1909	1.002	2.759
Head has any education	1058	0.120	0.325	851	0.170	0.376	1909	0.142	0.350
Head is female	1080	0.286	0.452	884	0.291	0.454	1964	0.288	0.453
Head is married	1080	0.821	0.383	884	0.872	0.334	1964	0.844	0.363
Head is widowed	1080	0.101	0.301	884	0.093	0.290	1964	0.097	0.296
Household size	1080	5.025	2.268	884	5.414	2.474	1964	5.200	2.370
Number of household members age 0-6	1080	1.213	1.226	884	1.324	1.364	1964	1.263	1.291
Number of household members age 7-15	1080	1.401	1.344	884	1.610	1.568	1964	1.495	1.452
Number of household members age 16-59	1080	2.321	1.176	884	2.317	1.275	1964	2.319	1.221
Number of household members age 60 and up	1080	0.090	0.302	884	0.164	0.411	1964	0.123	0.357

Source: Authors' computations using data form the PSNP4 End-line (2021) Household Survey.

Using these variables, we estimate propensity scores based on a logit model in which the dependent variable is the indicator for whether the household was a PSNP beneficiary for at least three of the last four years. As we did in the Midline Outcomes Report, we confirm balance by testing for equalityof means across the 2021 PSNP beneficiary and non-beneficiary samples as well as equality of the predicted propensity score across intervals or "blocks" of the predicted propensity score distribution. If balance is not achieved in the initial set of 5 intervals constructed, intervals leading to imbalance are divided into sub-intervals and the balance tests are re-run; having done so we assess common support using a standard common support graph.

2.4.4 Estimating the propensity score

Table 2.4 presents the results of the propensity score estimation based on a logit model for each of the three comparisons needed to construct the impact estimates. As noted by Imbens (2015, 389), "[…] the propensity score plays a mechanical role in balancing the covariates". In other words, the purpose of this exercise is to find a specification that leads to an accurate prediction of the program participation. Therefore, we do not spend time interpreting the regression coefficients.

Table 2.4: Propensity score regressions

	(1)	(2)	(3)
Treatment group:	*PSNP beneficiary in 2021*	*PSNP beneficiary in 2021*	*PSNP beneficiary in 2021*
Control group:	*PSNP non-beneficiary in 2021*	*PSNP beneficiary in 2016*	*PSNP non-beneficiary in 2016*
head age in years	-0.007	0.448	0.263
	(0.08)	(1.50)	(1.28)
head's age ^2	0.001	-0.007	-0.004
	(0.59)	(1.07)	(1.01)
head age in years ^3	-0.000	0.000	0.000
	(0.64)	(0.54)	(0.75)
education of hh head (in years)	0.151	-0.728	-0.172
	(0.81)	(1.05)	(0.33)
head's education ^2	-0.045	0.147	0.033
	(1.04)	(0.85)	(0.26)
head's education ^3	0.003	-0.007	-0.002
	(1.23)	(0.69)	(0.24)
household size is 2 or less	-0.820	0.434	-1.077
	(3.12)**	(0.44)	(1.67)
household size is 3 to 4	-0.606	0.662	-0.312
	(2.70)**	(0.92)	(0.56)
household size is 5 to 6	-0.301	0.940	0.165
	(1.40)	(1.31)	(0.31)
household size is 7 to 8	-0.091	1.744	0.869
	(0.41)	(2.20)*	(1.48)
number of old members (60+ years) / household size	0.179	5.514	0.127

	(0.31)	(2.30)*	(0.09)
HH has <0.5 ha of land	0.259	2.309	4.784
	(1.31)	(2.23)*	(5.83)**
HH has more than 1 ha of land	0.086	-4.049	-0.984
	(0.38)	(3.74)**	(1.29)
HH has 1 or more oxen	-0.235	-0.489	-2.377
	(0.83)	(0.44)	(2.81)**
HH has 3 or more heads of cattle	0.083	-1.087	0.221
	(0.37)	(1.46)	(0.44)
HH has 1 or 2 head of cattle	-0.000	-0.725	1.228
	(0.00)	(0.78)	(1.76)
HH has 5 or more sheep/goats	0.193	2.434	1.797
	(0.97)	(4.11)**	(4.01)**
HH has 5 or less chicken & no other livestock	-0.092	2.526	3.959
	(0.31)	(2.10)*	(2.36)*
HH owns no animals	0.399	0.164	-0.564
	(1.25)	(0.14)	(0.34)
Household has a corrugated iron roof	-0.473	0.508	-1.475
	(2.01)*	(0.79)	(3.56)**
Household's house is in bad shape	0.059	-0.553	0.292
	(0.39)	(0.82)	(0.56)
=household owns a mobile phone	0.138	-0.788	0.909
	(0.96)	(1.35)	(1.98)*
household received 100 birr or more in remittances per month	0.499	1.299	0.125
	(1.04)	(0.14)	(0.11)
tropical livestock units (TLU) owned by hh	0.027	-0.007	-0.060
	(2.35)*	(0.16)	(2.83)**
an adult member in HH cannot work because of injury or poor	0.103	0.643	0.808
hearing/sight	(0.37)	(0.78)	(1.23)
A little poorer than most households	0.385	1.100	0.900
	(2.62)**	(1.89)	(2.17)*
Amongst the poorest in the village	0.459	1.322	0.670
	(3.07)**	(2.30)*	(1.72)
The poorest in the village	0.199	1.091	1.475
	(1.05)	(1.48)	(2.39)*
Household thinks it's same as 2 years ago	0.247	1.109	0.713
	(1.74)	(2.07)*	(1.89)
Household thinks it's worse off than 2 years ago	-0.045	0.673	1.115
	(0.29)	(1.22)	(2.52)*
HH has less than 0.5 ha of land and is considered eligible for PSNP	-0.125	0.304	-1.619
	(0.64)	(0.46)	(2.66)**
Interaction: land1_psnp_eXR_2	0.346	-8.305	-7.695
	(1.09)	(5.03)**	(7.81)**
HH has between 0.5 and 1 ha of land and is considered eligible for	0.653	6.110	22.186
PSNP	(2.00)*	(3.99)**	(20.17)**
Interaction: land2_psnp_eXR_2	0.181	-16.080	-25.884
	(0.33)	(7.06)**	
HH has more than 1 ha of land and is considered eligible for PSNP	-0.699	7.666	4.497
	(0.82)	(2.10)*	(2.45)*
Interaction: land3_psnp_eXR_2	1.103	-12.424	-7.139
	(1.19)	(3.08)**	(3.62)**
HH has one or more oxen and is considered	0.701	0.385	2.080

eligible for PSNP			
	(1.13)	(0.18)	(1.66)
Interaction: ls1_psnp_eXR_2	-0.089	-9.872	-1.284
	(0.09)	(0.00)	(0.61)
HH has three or more heads of cattle and is considered eligible for	-0.069	1.386	1.399
PSNP	(0.15)	(1.07)	(1.76)
Interaction: ls2_psnp_eXR_2	0.663	26.031	-0.712
	(0.97)		(0.33)
HH has one or two heads of cattle and is considered eligible for PSNP	-0.803	2.132	-1.200
	(1.57)	(0.93)	(1.04)
Interaction: ls3_psnp_eXR_2	0.657	26.008	0.977
	(1.16)	(0.01)	(0.71)
HH has five or more sheep/goats and is considered eligible for PSNP	0.120	-3.202	-3.026
	(0.61)	(3.62)**	(4.86)**
Interaction: ls4_psnp_eXR_2	-0.385	4.993	3.270
	(1.32)	(3.92)**	(3.62)**
HH has a few chickens but no other animals and is considered eligible	-0.127	-1.722	-2.601
for PSNP	(0.37)	(1.18)	(1.57)
Interaction: ls5_psnp_eXR_2	-0.492	-2.322	-3.449
	(0.86)	(1.08)	(2.17)*
HH has no animals and is considered eligible for PSNP	-0.615	-0.995	-1.988
	(1.73)	(0.69)	(1.18)
Interaction: ls6_psnp_eXR_2	-0.100	5.784	2.596
	(0.19)	(2.00)*	(1.57)
HH's dwelling has metal roofing and is considered eligible for PSNP	-0.081	22.107	22.785
	(0.25)	(8.53)**	(16.05)**
Interaction: oas1_psnp_eXR_2	0.626	-22.002	-21.379
	(1.38)		
HH's dwelling is obviously in a state of disrepair and is considered	-0.179	1.724	0.435
eligible for PSNP	(0.72)	(1.70)	(0.65)
Interaction: oas2_psnp_eXR_2	-0.403	-7.809	-5.869
	(0.77)	(4.28)**	(5.52)**
HH owns a mobile phone and is considered eligible for PSNP	-0.104	1.752	-0.692
	(0.51)	(2.23)*	(1.21)
HH receives 100 birr remittances each month and is considered	0.347	-6.756	-3.434
eligible for PSNP	(0.47)	(0.73)	(1.77)
household experienced a non-drought shock in the last 2 years	0.107	1.207	0.380
	(0.87)	(2.45)*	(1.20)
kebele did not have sufficient funds for all eligible households	0.046	3.964	1.225
	(0.30)	(4.82)**	(3.20)**
Community is connected to a road made of stones	0.199	-17.444	-9.139
	(0.59)	(7.61)**	(5.61)**
Community is connected to a dirt road	-0.058	-7.147	-4.352
	(0.36)	(8.57)**	(10.22)**
community received humanitarian food aid	0.237	-53.255	-29.615

	(1.42)	(36.25)**	
kebele officials received training on targeting in the last 12 m	1.032	-6.393	-1.827
	(5.38)**	(7.49)**	(5.06)**
Only the poorest households in this woreda have access to PSNP benefits	-0.918	-1.623	1.133
	(5.41)**	(2.61)**	(2.69)**
region==Afar	0.710	42.964	22.414
	(2.73)**		(25.08)**
Interaction: oinc1_psnp_eXR_2		18.648	14.260
		(0.01)	
Constant	-1.314	1.761	-1.086
	(1.02)	(0.37)	(0.33)
N	1,716	1,576	1,582

Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Surveys.
Note: * $p<0.05$; ** $p<0.01$

2.4.5 Common support in the predicted propensity score

Figure 2.1 below shows the standard common support graph for the three cases. We see that the region of common support is wide and dense in the first model (Figure 2.1a) of participation comparing PSNP beneficiaries to non-beneficiaries in 2021. The region of common support is significantly more limited for the second and third models (Figures 2.1b and 2.1c), comparing participation of PSNP beneficiaries in 2021 to beneficiaries in 2016 and comparing participation of PSNP beneficiaries in 2021 to non-beneficiaries in 2016, respectively. While the cross-sectional matching model often faces challenges with finding strong common support across all three models (so this weakness in models is not uncommon), we will explore alternative specifications with stronger common support.

Figure 2.1: Common Support Graphs

Figure 2.1a: *PSNP beneficiary in 2021 vs PSNP non-beneficiary in 2021*

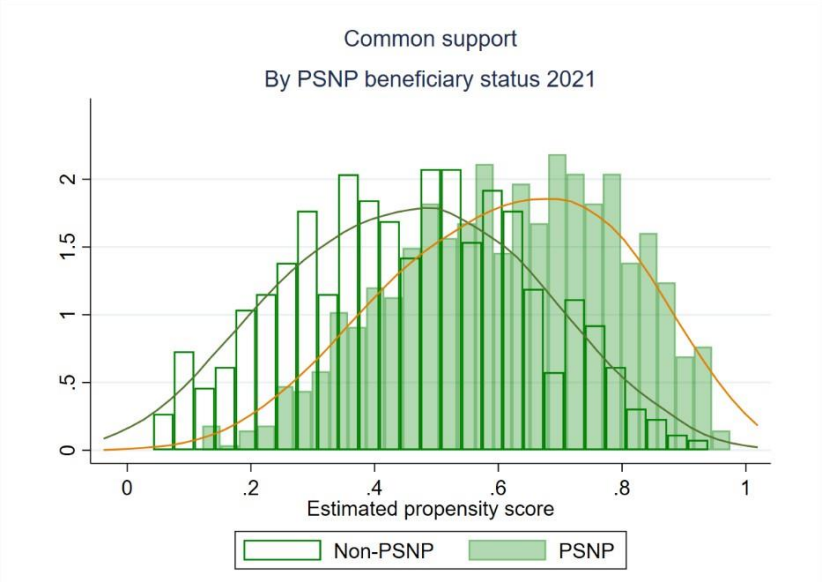


Figure 2.1b: *PSNP beneficiary in 2021 vs PSNP beneficiary in 2016*

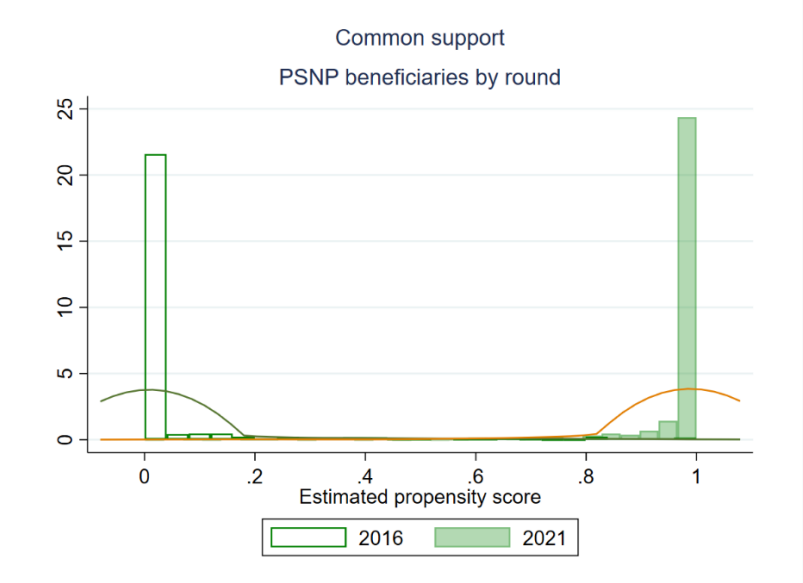
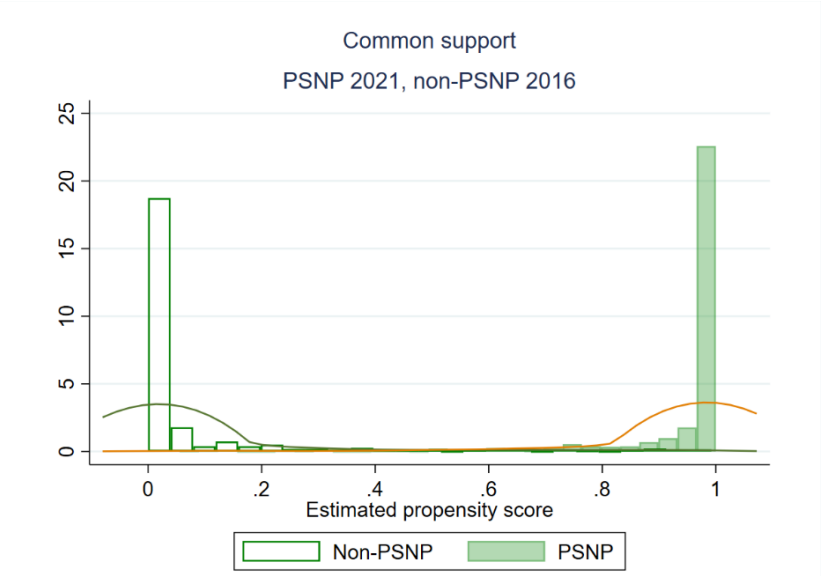


Figure 2.1c: *PSNP beneficiary in 2021 vs Non-PSNP beneficiary in 2016*



Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Surveys.

2.4.6 The choice of matching algorithm

Blundell and Dias (2009) propose using kernel matching algorithm in the context of repeated cross-section and difference-in-difference matching. In our case, kernel matching estimator uses weighted averages of all households within neighbourhood in the three control groups to construct the counterfactual outcome. Using all observations, instead of only the nearest ones, the kernel matching has the advantage of reducing variability of the estimator. The drawback is that the analytical standard errors will not be valid. Instead, the standard errors in this case need to be computed using bootstrap methods.

Chapter 2 references

Abadie, A and G. Imbens. 2006. Large Sample Properties of Matching Estimators for Average Treatment Effects *Econometrica* 74(1): 235-267.

Berhane, Guush, Daniel O. Gilligan, John Hoddinott, Neha Kumar and Alemayehu Seyoum 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.

Blundell, R. and M.C. Dias. 2009. Alternative approaches to evaluation in empirical microeconomics. Journal of Human Resources 44(3):565-640.

Richard Blundell, Monica Costa Dias, Costas Meghir, John van Reenen, Evaluating the Employment Impact of a Mandatory Job Search Program, Journal of the European Economic Association, Volume 2, Issue 4, 1 June 2004, Pages 569–606.

Gilligan, Daniel O., John Hoddinott and Alemayehu Seyoum Taffesse. 2009. "The Impact of Ethiopia's Productive Safety Net Program and Its Linkages." Journal of Development Studies 45(10): 1684-1706.

Heckman, J. J., H. Ichimura, and P. E. Todd. 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. Review of Economic Studies 64: 605-654.

_____. 1998. Matching as an econometric evaluation estimator. Review of Economic Studies 65: 261-294.

Hirvonen, Kalle, Alemayehu Seyoum Taffesse, and Ibrahim Worku. 2016. "Seasonality and household diets in Ethiopia." Public Health Nutrition 19 (10):1723-1730.

Imbens, Guido W. 2015. "Matching methods in practice: Three examples." Journal of Human Resources 50 (2):373-419.

Kazianga, Harounan, and Christopher Udry. 2006. "Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso." Journal of Development Economics 79 (2):413-446.

Rosenbaum, P., and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1): 41-55.

Smith, Jeffrey & E. Todd, Petra, 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?," Journal of Econometrics, vol. 125(1-2):305-353.

Chapter 2 Annex: The role of seasonality

Due to unforeseen circumstances, mostly around the COVID-19 pandemic, the midline and end-line survey timing were delayed. The midline survey took place approximately 6 months later than initially planned and the end-line survey took place 10 months later in the calendar. As a result, the baseline and the midline survey rounds took place in different seasons.[3] The baseline data were collected in January/February 2016, which, for most areas, is a postharvest season characterized by dry conditions. In contrast, the midline survey took place in June/July 2018, which for most part of the country is a rainy and lean season. The end-line survey took place in April/May 2021, prior to the rainy season. These seasonal differences in the timing of the outcome variable measurement can affect the estimated size of the effect of the program for variables that are more affected by seasonality. This should not introduce bias into the impact estimates, however, because both PSNP and non-PSNP households experience the same shift in seasons and this change in seasons is not caused by PSNP participation.

Some of the PSNP outcome indicators are more sensitive than others to the season in which the data are collected. We believe that the food gap is perhaps among the least sensitive to these changes because the recall period is 12 months. In contrast, household food consumption and dietary diversity indicators are based on 7-day recall and based on our previous work subject to considerable seasonal fluctuations, especially in rural areas (Hirvonen, Taffesse, and Worku 2016). The indicators measuring asset portfolios, especially livestock, may also fluctuate across seasons. However, based on the available evidence from other countries (Kazianga and Udry 2006), compared to consumption, we should expect livestock holdings to remain relatively constant across seasons.

[3] The midline survey was not used in the matching analysis.

Chapter 3: How Do Beneficiaries Use Their Transfers?

3.1 Introduction

In this chapter, we discuss how PSNP beneficiary households in Afar and Somali regions used their last cash and food transfers. The chapter will address **RQ17: Use of transfers**. Section 3.2 discusses the proportion of households that spent their cash transfers on different expenditure categories and the importance of these categories in total cash transfers. The survey instrument identifies ten expenditure categories for cash transfers. In the description below we aggregate these into six categories. The three types of non-food expenditures (those that directly benefit children, male adults, and female adults) are aggregated into one category, non-food expenditures. Cash transfers that recipients were asked to give to other households by anyone in a position of authority is combined with cash transfers used to voluntarily help other households to form a category labelled as "Other households". Finally, cash transfers given to local authorities is aggregated with cash transfers given to persons designated to collect PSNP payments into a category labelled as "Others". It is important to note that fewer than 10 households mentioned that they gave some of their cash/in-kind (food) transfer to a person in a position of authority or were asked by such a person to give some of their transfer to other households.

Section 3.3 describes the proportion of households that used their last food transfers into one or more use categories food transfers can be put into. We also describe the importance of these categories in total food transfers. In the descriptions, the six food-use categories identified in the survey instrument are aggregated into four and these involve the last two aggregations described above for cash transfers.

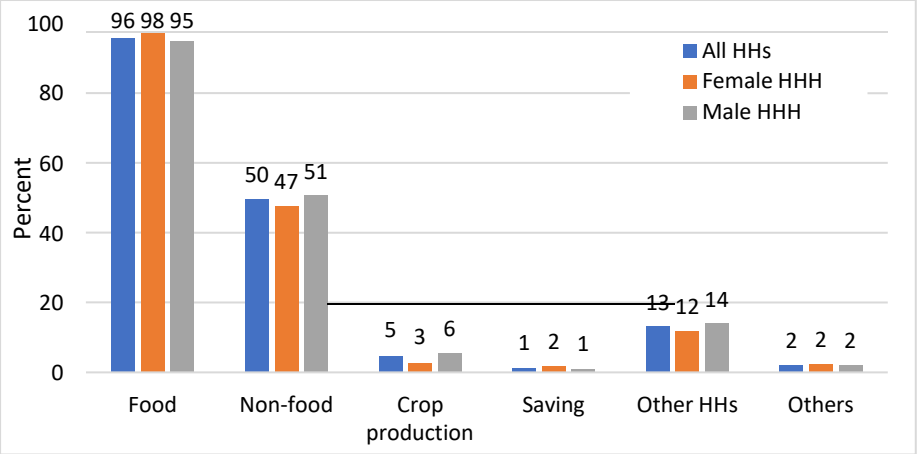
3.2 Cash transfers

This section first describes the proportion of households that use their cash transfers for different purposes, which are summarized in Figures 3.1 and 3.2, and the first panel in Table 3.1. It describes how many (what percentage) of the households used their cash transfers to buy food or non-food items. Typically, the sums of the percentages of households that used cash transfers on different expenditure categories adds up to higher than a 100 percent because each household can spend cash transfers on each expenditure category. Then, the section discusses the importance of each expenditure category in total cash transfers of the household. The sum of the percentages of cash transfers spent on all categories sums up to a 100 percent. These are summarized in Figures 3.3-3.5 and the second panel in Table 3.1.

Figure 3.1 shows the proportion of households that used cash transfers for different purposes by gender of household head. About 96 percent of the households used their cash transfers to buy food. Nearly 50 percent used cash transfers to buy non-food items and 13 percent shared them with other households. Relative to male headed households, the proportion of female headed households that purchased food is slightly but statistically significantly higher. Female headed households that purchased non-food items, inputs used in crop production, and that shared with other households is

slightly lower than the corresponding proportion of male headed households, although these differences are not statistically significant.

Figure 3.1: Households that used cash transfer for different purposes (%), by gendered household type

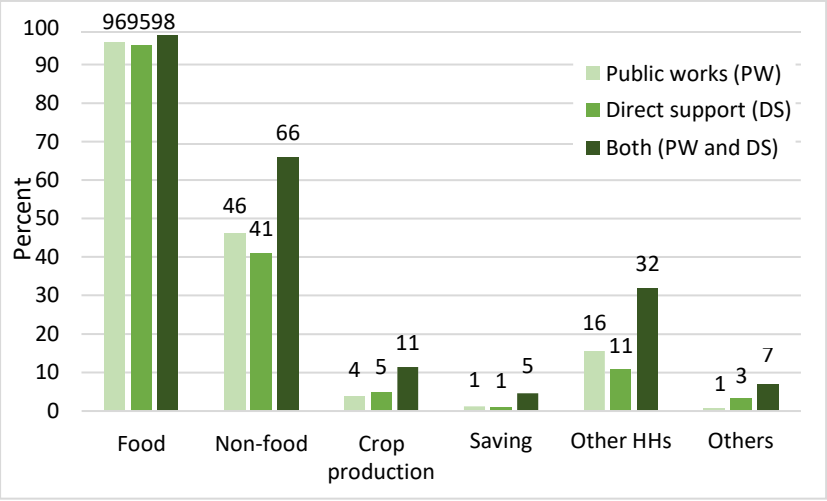


Note: HHs = households; HHH = Headed household.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figure 3.2 shows the proportion of households that used cash transfers for different purposes across PSNP status. The figure indicates that a higher proportion of households that participate in both public works and direct support use cash transfers to purchase non-food items and to support other households. A lower proportion of households on direct support purchase non-food items and share their cash transfers with other households. The differences indicated are statistically significantly different from zero.

Figure 3.2: Households that used cash transfer for different purposes (%), by PSNP status



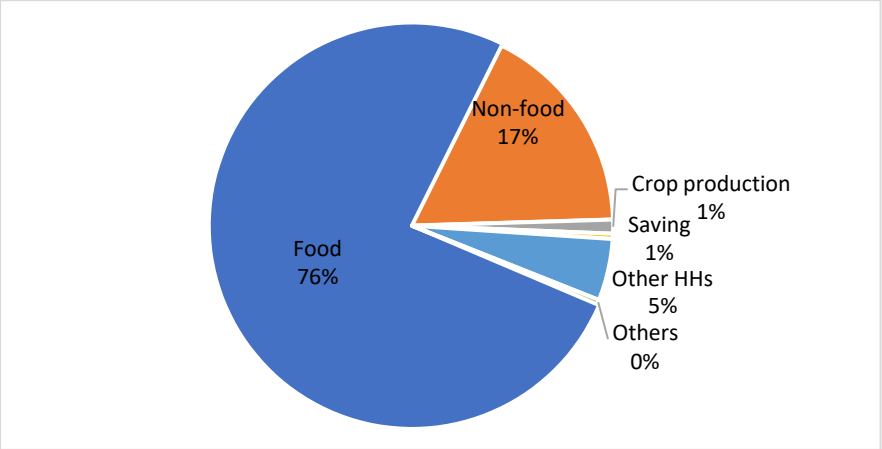
Note: HHs = households.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figure 3.3 summarizes the share of different expenditure categories in total cash transfer for all households. The figure indicates that about three-quarters of the cash transfers were used on food

purchases. About 17 percent of the cash transfers were used to purchase non-food items, out of which 11 percent was spent on non-food goods that directly benefit children and a slightly higher (3.5 percent) share of the remaining on non-food items that directly benefit male adults. Nearly 5 percent of the cash transfers was used to help other households.

Figure 3.3: Percentage of cash transfer used for different purposes

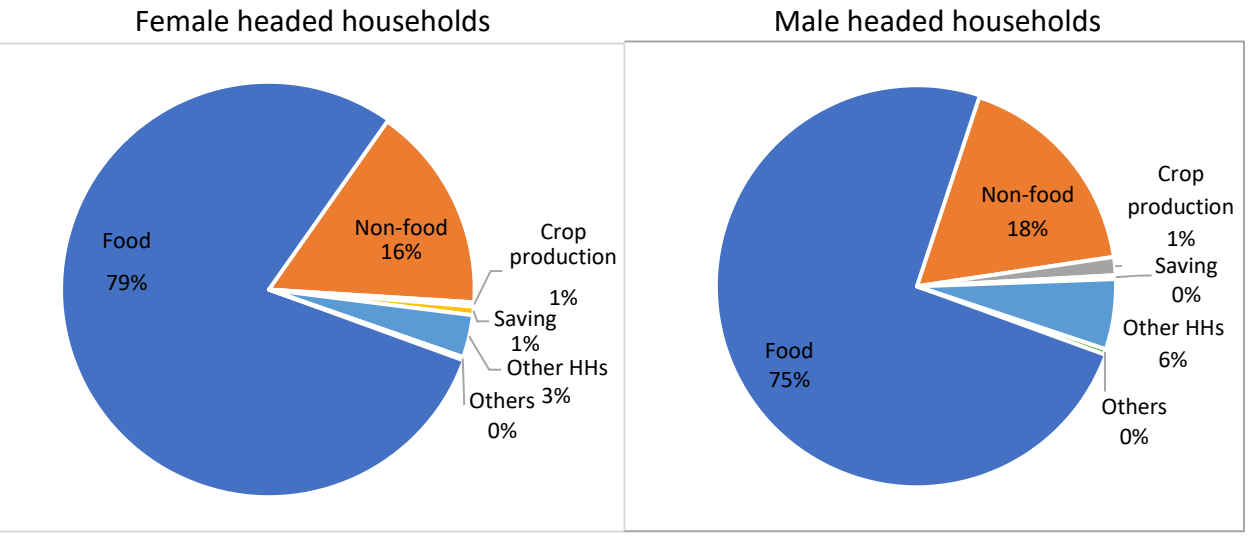


Note: HHs = households.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figures 3.4 below shows the importance of different expenditure categories across gendered household types. Female headed households spent a slightly higher share of their cash transfers on food, and a slightly lower share on non-food items and to help other households. Except for the difference in the share spent on the purchase of non-food items, the differences stated are statistically significant.

Figure 3.4: Percentage of cash transfer used for different purposes, by gendered household type



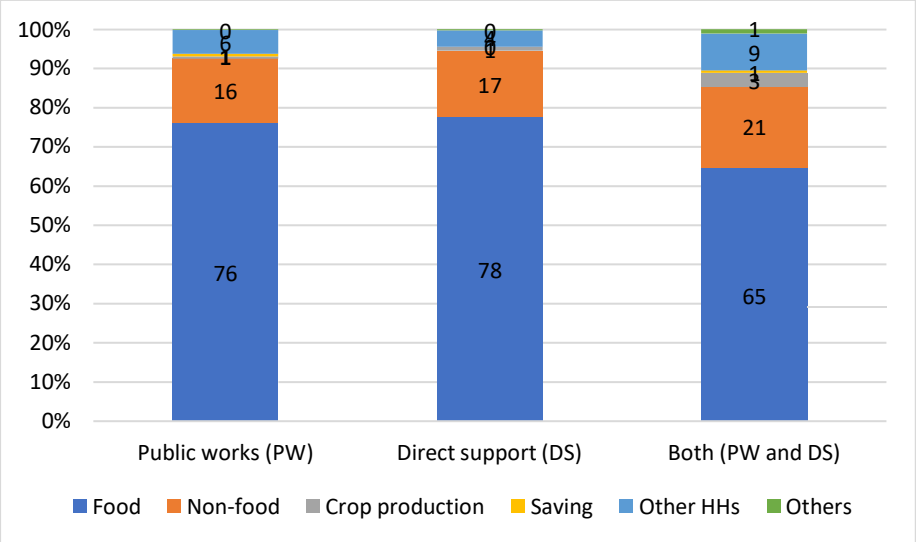
Note: HHs = households.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figure 3.5 indicates that households that participate in both public works and direct support spent a relatively lower share of their cash transfers to buy food, but a higher share to buy non-food items

and, inputs for crop production. The share of their support to other households is also greater. However, out of these it is only the difference in the share spent to buy food that is statistically significant.

Figure 3.5: Percentage of cash transfer used for different purposes, by PSNP status



Note: HHs = households.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

In Table 3.1 we summarize the data on the proportion of households that used cash transfers for different purposes and the importance of these purposes in total cash transfers across regions. The table indicates that a higher proportion of households in Somali region spent cash transfers on all expenditure categories and the differences are statistically significant for all expenditure categories except other expenditures. However, there was little difference in the importance of different expenditure categories in cash transfers of households in Afar and Somali regions. Consequently, all differences in the shares of expenditures except saving are not statistically significant.

Table 3.1: Households using cash transfers and proportion of cash transfers used for different purposes (%), by region

	Households that used cash transfer (%)		Cash transfers used (%)	
	Afar	Somali	Afar	Somali
Food	91.5	97.2	77.8	75.4
Non-food	42.6	52.0	17.7	17.0
Crop production	0.7	6.0	0.7	1.2
Saving	0.0	1.6	0.0	0.6
Other HHs	5.0	16.2	3.2	5.5
Others	1.4	2.3	0.6	0.3

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

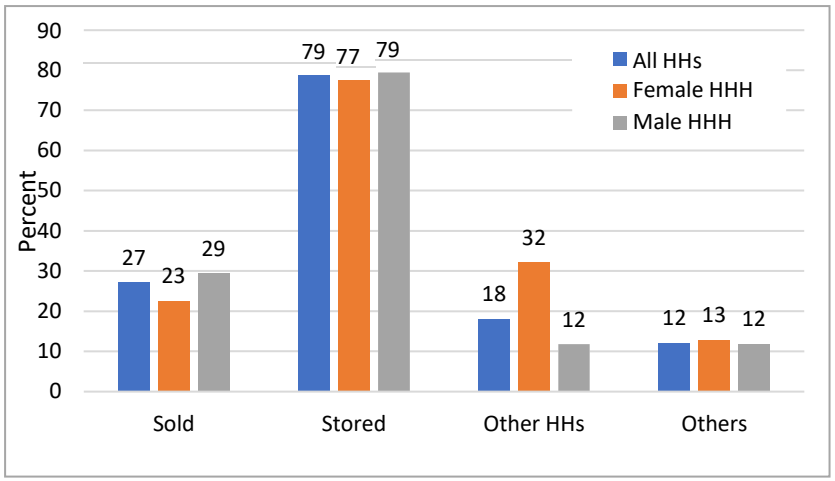
3.3 Food transfers

This section first describes the proportion of households that use their food transfers for different purposes. For instance, it describes the percentage of households that sold food transfers or those

that have food transfers in storage. Since all or most households may have used food transfers for most purposes the sums of these percentages may add up to higher than a 100 percent. These are summarized in Figures 3.6 and 3.7, and the first panel in Table 3.2. Then, the section discusses the importance of each use category in total food transfers of the household. The sum of the shares of all use categories adds up to a 100 percent. These are summarized in Figures 3.8-3.10 and the second panel in Table 3.2.

Figure 3.6 shows that 22 percent of households that received food transfers sold food while nearly 80 percent have food in storage at the time of the survey. Almost all 12 percent of the households that put food transfers to other uses gave it to persons designated to collect PSNP payments. Female headed households that gave food transfers for other households (32 percent) is over 20 percent higher than the proportion of male headed households that used food transfers for the same purpose and this difference is statistically significant.

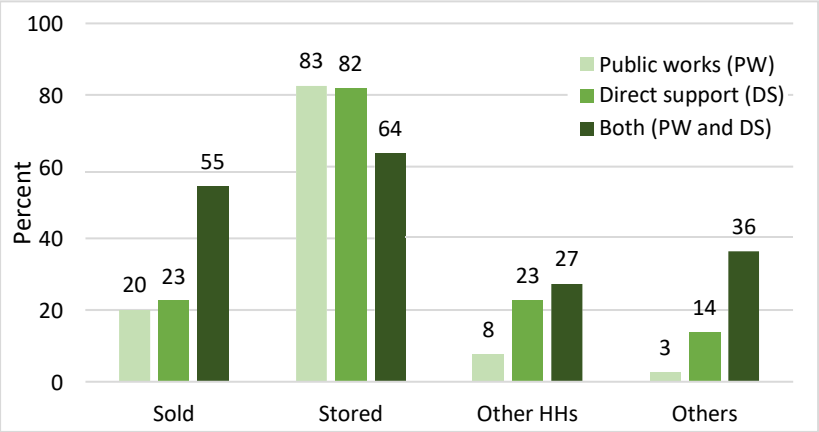
Figure 3.6: Households that used food transfers for different purposes (%), by gendered household type



Note: HHs = households; HHH = Headed household.
Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

A relatively lower proportion of households that participated in both public works and direct support have food in their storage while a relatively higher proportion of these households used food transfers for the remaining three purposes. The reverse holds for households that participated in only public works. However, the only statistically significant difference is the proportion of households that used food transfers for other purposes.

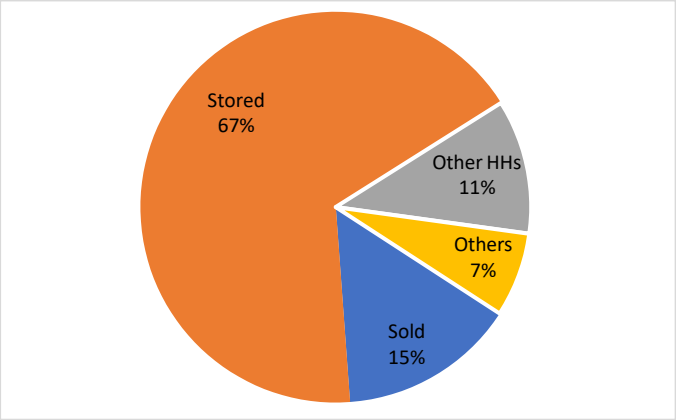
Figure 3.7: Households that used food transfers for different purposes (%), by PSNP status



Note: HHs = households.
Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Most of the food transfers were in storage at the time of the survey (Figure 3.8). The share of food transfers given to other households (11 percent) is considerably higher than the share of cash transfers given to other households (5 percent), indicating that households are more likely to provide in-kind assistance for other households. Moreover, all of the food given to other households was meant to help those households (and not because anyone in a position of authority told the households to share the transfers to other households).

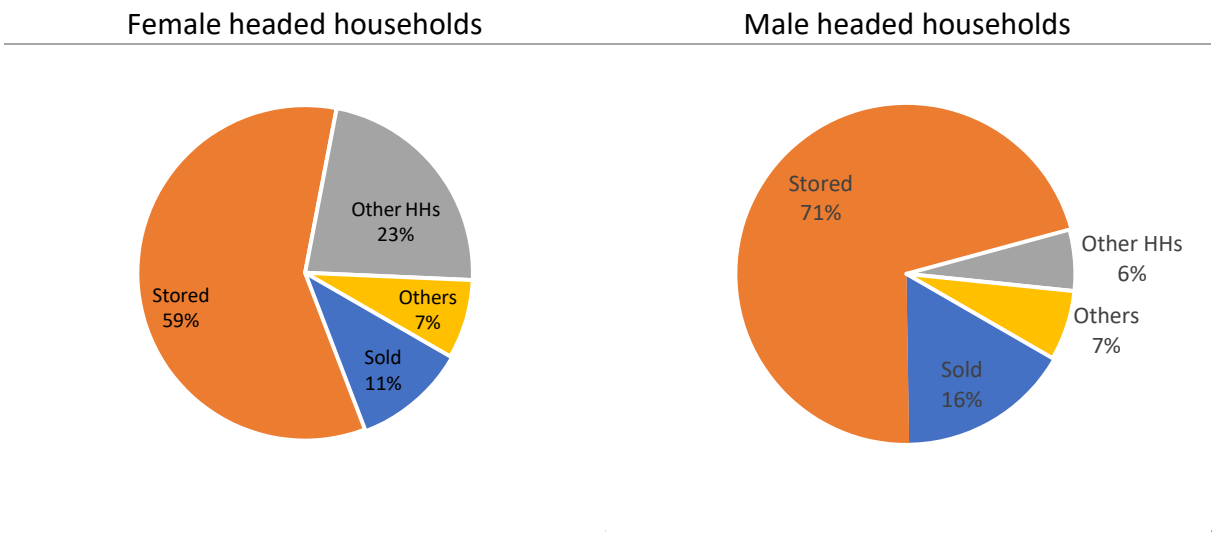
Figure 3.8: Percentage of food transfers used for different purposes,



Note: HHs = households.
Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Relative to male headed households, female headed households give a considerably higher share of their food transfers for other households (Figure 3.9) and the difference is statistically significant. Female headed households have a considerably lower share in storage and sold a slightly lower share of their food transfers, although both of these differences are not statistically significant.

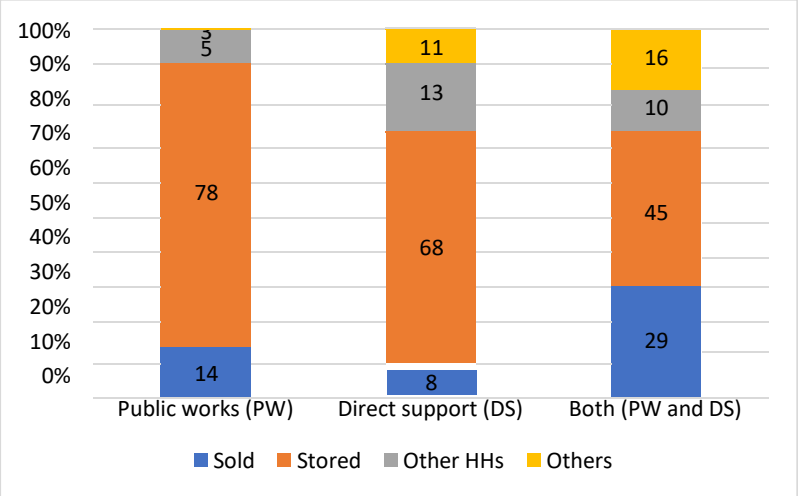
Figure 3.9: Percentage of food transfers used for different purposes, by gendered household type



Note: HHs = households.
Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Households participating in both public works and direct support sold 29 percent of their food transfers. Although this is over twice the next higher share by households participating in public works (14 percent), the difference is not statistically significant. In contrast, the proportion of food in storage by households participating in both public works and direct support is nearly 20 percent less than the next lower share and this difference is statistically significant. Relative to other households those participating in public works used a considerably lower share of their food transfers to help other households and for other purposes although differences in shares of both use categories are not statistically significant.

Figure 3.10: Percentage of food transfers used for different purposes, by PSNP status



Note: HHs = households.
Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Table 3.2 summarizes the data on how food transfers are used across regions. Relative to households in Afar, a considerably and statistically significantly higher proportion of those in Somali region used

food transfers on all use categories. This excludes a slightly lower and statistically not significant

proportion of households in Somali region that stored food. Consistent with this, households in Somali region sold, shared with other households, and used for other purposes a considerably higher proportion of their food transfers. In contrast, the share of food stored by households in Somali region (20 percent) is less than a third of the corresponding share stored by households in Afar (71 percent). Except for the proportion of food that was sold all of the differences noted are statistically significantly different from zero.

Table 3.2: Households using food transfers and proportion of food transfers used for different purposes (%), by region

	Households that used food transfer (%)		Food transfer used (%)	
	Afar	Somali	Afar	Somali
Sold	22.0	87.5	14.0	22.9
Stored	79.1	75.0	71.3	20.2
Other HHs	12.1	87.5	9.1	34.3
Others	6.6	75.0	5.6	22.7

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

3.4 Summary

The results indicate that nearly all households in lowland areas spend cash transfers to buy food and 50 percent to buy non-food items. Out of the remaining proportions the only one that exceeds five percent is the proportion of households that use cash transfers to help other households (13 percent). Households in lowland areas use over three-quarters of cash transfers to buy food and about 17 percent to buy non-food items. The share of cash transfers given for other households is five percent while the share used for the remaining purposes is one percent or lower. Most of the food transfer is in storage, nearly 15 percent has been sold, and 11 percent was given to other households.

Chapter 4: Food Security and Household Diet Diversity

4.1 Introduction

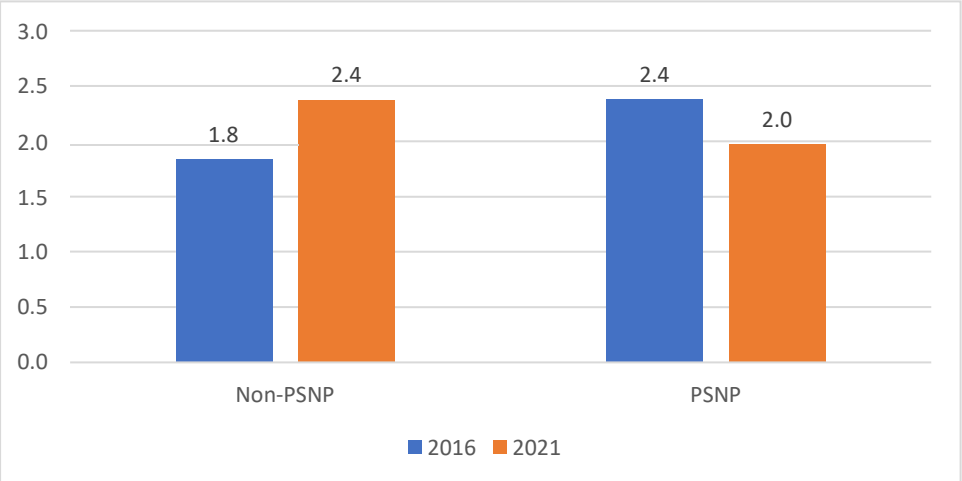
Food security and dietary diversity are two important outcome indicators considered in the evaluation of the PSNP4. This chapter does two things. First, it reports descriptive results of trends in food security and household level dietary diversity, comparing mean outcomes between baseline and end-line. Second, it summarizes the impact of the PSNP on these two outcomes.

As in previous rounds, food insecurity is measured by food gap, a household level food insecurity measure that summarizes self-reported number of months the household has been unable to satisfy its food needs in the 12 months preceding the survey month. Household dietary diversity score (HDDS) is a score that measures the number of different food groups consumed over a given reference period. In sections 4.2 and 4.3, we report the descriptive results showing trends between baseline and end-line. We also disaggregate some of these results by PSNP status, gender of the household head, and region. In section 4.4, we report the summary of the impact estimates and section 4.5 summarizes the chapter. We note a caveat that the Lowland data is a cross-sectional time series and not panel data. The implication is that unlike in the Highlands, here a new sample is selected in each round and the mean outcomes represent mean differences of outcomes between the two cross-sections, and not same households (see also Chapter 2).

4.2 Trends in food security

As described above, food gap is calculated by taking the average of the responses to the question: "How many months in the last 12 months did you have problems satisfying the food needs of your household?". The same question was administered at baseline and end-line for both PSNP and non-PSNP households. This measure ranges between zero (no month of food insecurity) and 12 (full-year food insecurity), and the mean value is interpreted as the number of months of food insecurity faced by the average household. Looking at the trends, a reduction (increase) in average food gap between two periods means an improvement (deterioration) in food security.

Figure 4.1: Mean food gap by PSNP and survey round

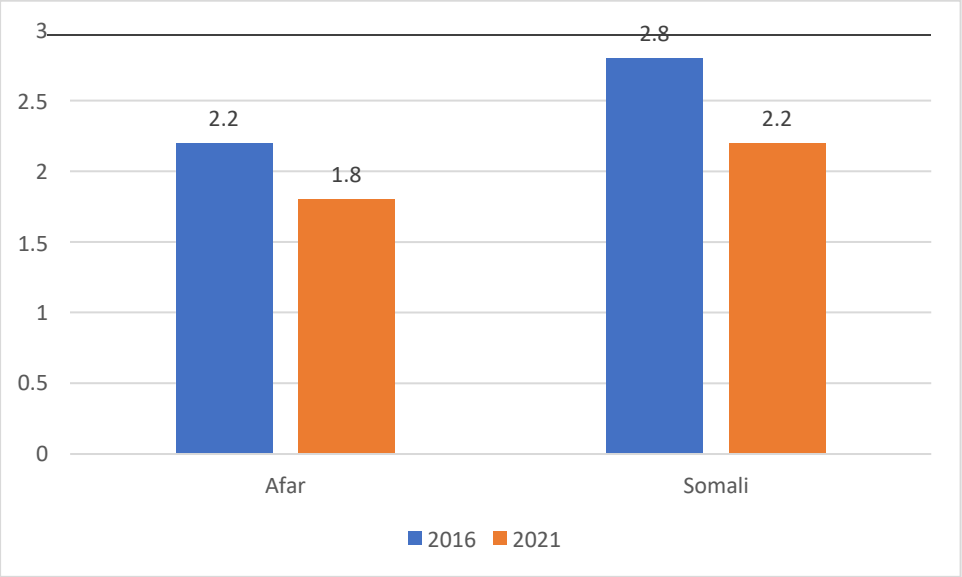


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

We begin by presenting overall trends in mean food gap for the lowlands, by PSNP status. Figure 4.1 shows trends in mean food gap for the two regions between baseline and end-line, disaggregated by PSNP beneficiary status. The striking trend is that between 2016 and 2021, food gap has increased on average for non-PSNP beneficiaries but decreased for the PSNP beneficiaries. That is, mean food gap between baseline and end-line increases by 0.6 months (from 1.8 months in 2016 to 2.4 months in 2021) for the non-PSNP beneficiaries and decreased by 0.4 months (from 2.4 months in 2016 to 2.0 in 2021) for the PSNP households. We note that unlike in the Highlands, the baseline and end-line Lowland samples are two different cross-sectional samples. A simple mean t-test shows these differences are statistically significant at the 95 percent significance level. It is important to note however that this Lowland-level average masks a lot of heterogeneities.

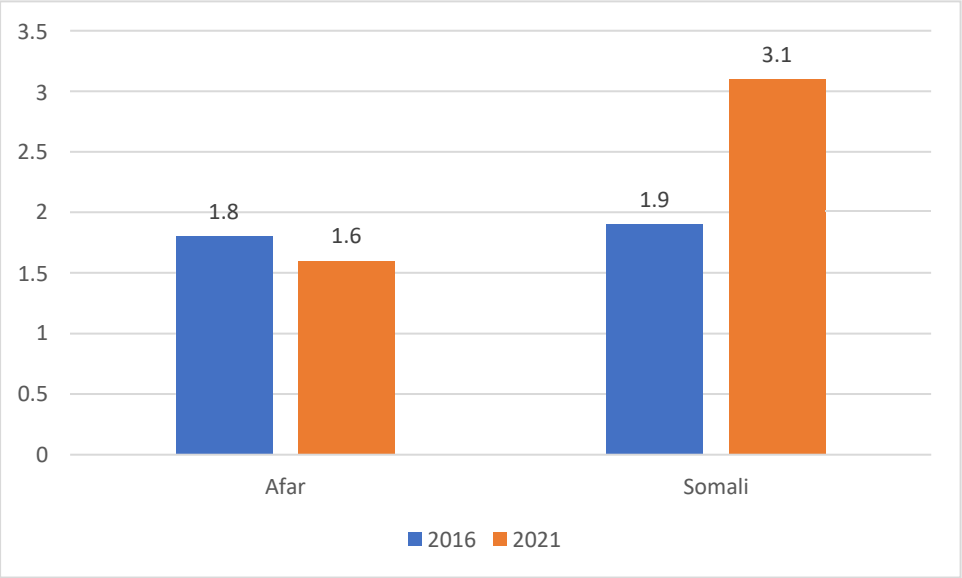
Further disaggregating these by PSNP beneficiary status and survey round reveals a different picture for each region. Figure 4.2 shows these for (a) PSNP beneficiaries and (b) PSNP non-beneficiaries, by region and survey round. And, we see that for the PSNP beneficiaries in both regions, mean food gap has declined – by 0.4 month in Afar and 0.6 month in Somali (Figure 4.2 (a)). Mean food gap has also declined among non-PSNP households in Afar (by 0.2 month) but increased among non-PSNP households in Somali (by 1.2 months). Overall, we note that mean food gap is higher among households in Somali than among households in Afar, regardless of PSNP status. Note also that these trends are similar to those reported in the midline outcome reports (albeit the averages reported here are on the high side than those reported at midline).

Figure 4.2(a): Mean food gap among PSNP households by region and by survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Figure 4.2(b): Mean food gap among non-PSNP households by region and by survey round

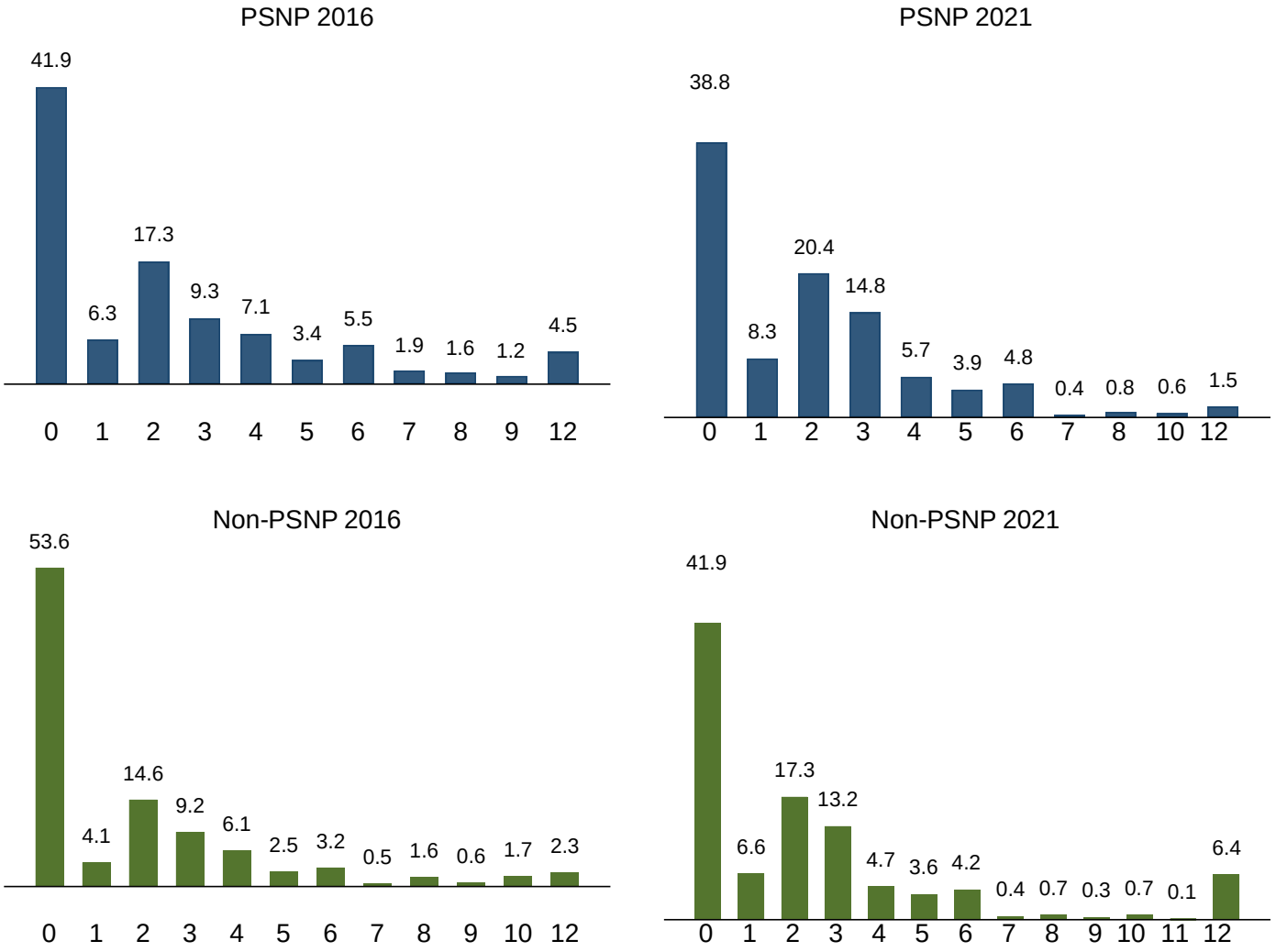


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

The mean food gap presented so far, while providing the average food gap picture among different groups, still hides an important heterogeneity within a specific group. As discussed earlier, our food gap measure comes from the survey question that asked about the number of months the household has been food insecure over the 12 months before the survey. This allows us to see what proportion of households are food secure or food insecure in each category of households. Figure 4.3 shows the distribution of food gap by survey round and PSNP beneficiary status.

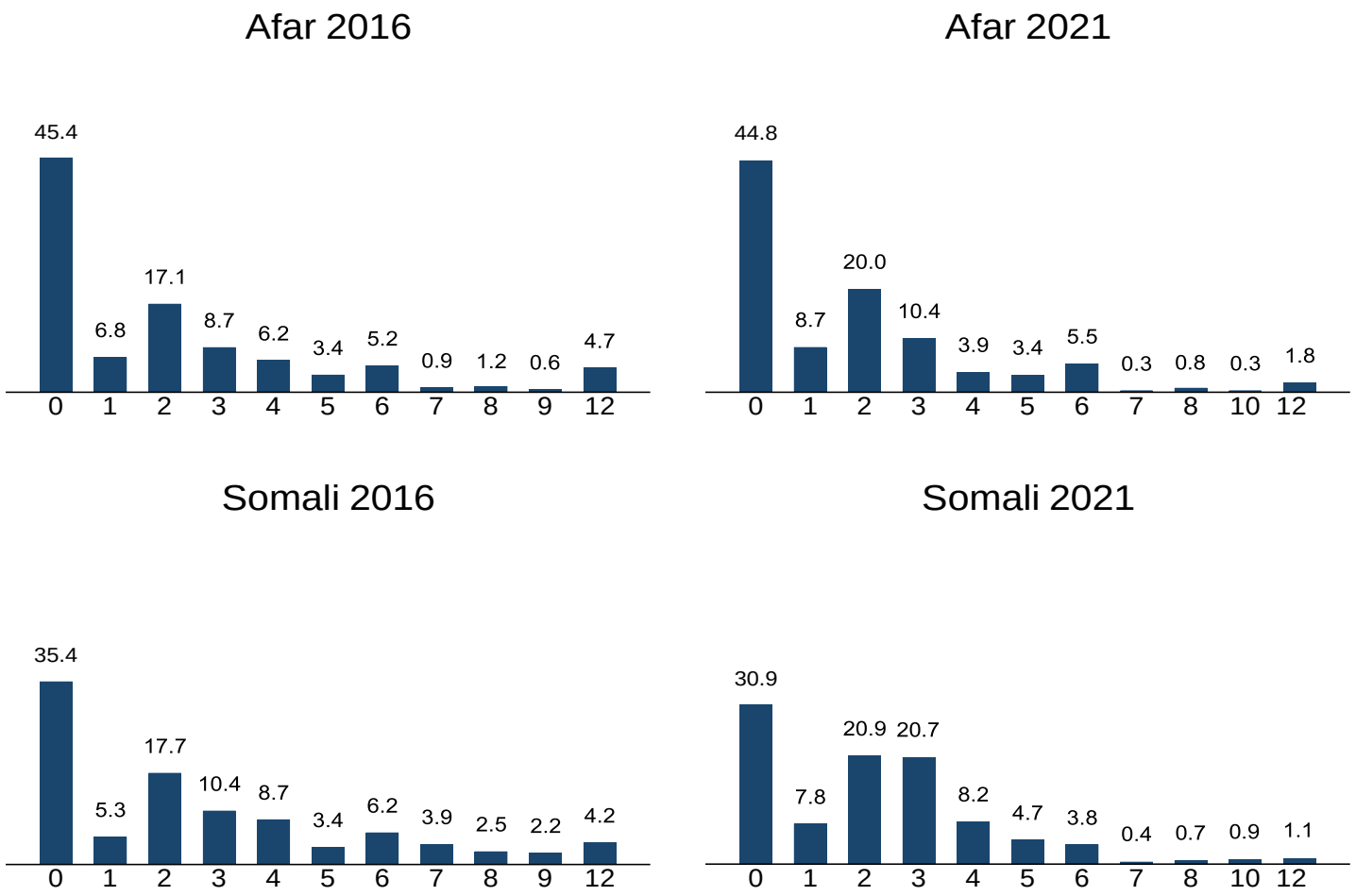
We note three important takeaways Figure 4.3. First, a large share of PSNP (e.g., 39 percent at end-line) and non-PSNP (e.g., 42 percent at end-line) households report that they did not have food insecurity problems – i.e., had zero food gap (Figure 4.3). Second, the share of PSNP and non-PSNP households reporting zero food gap have however declined between baseline and end-line – a decline by 3 percentage points among the PSNP and 11 percentage points among the non-PSNP households.

Figure 4.3: Distribution of food gap by PSNP beneficiary status and survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Figure 4.4: Distribution of food gap among PSNP households by survey round and region



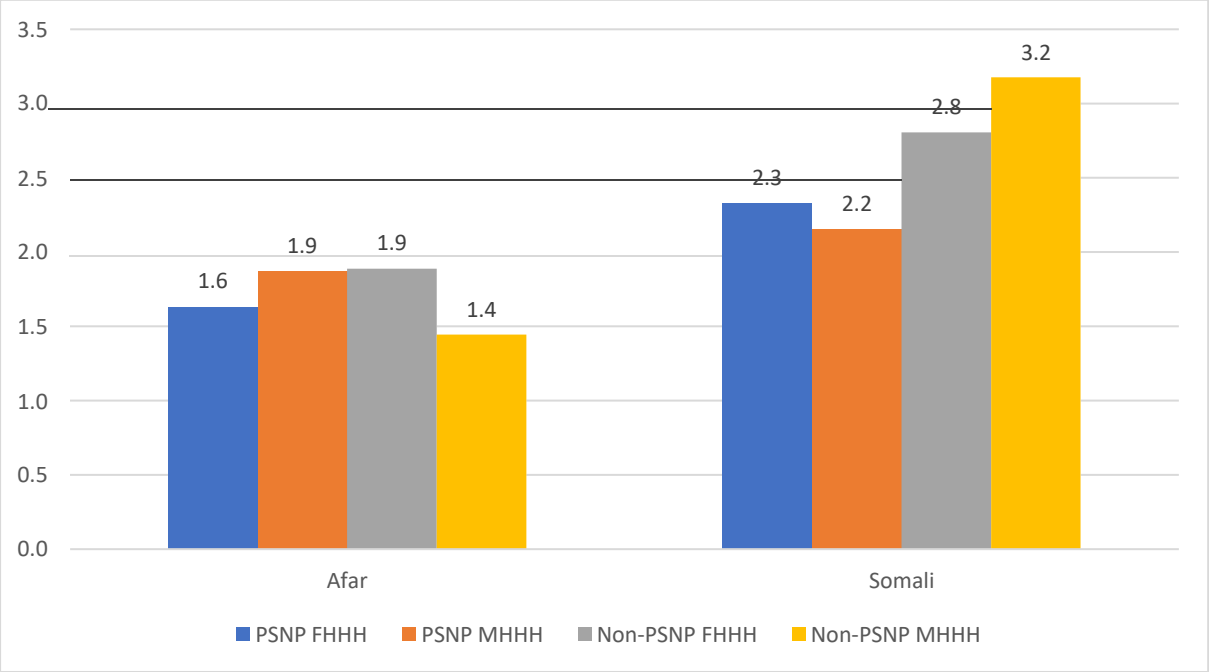
Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Third, between baseline and end-line, the share of households reporting a food gap of three months and above has declined by 2 percentage points (from 35 percent in 2016 to 33 percent in 2021) for the PSNP but increased by 6 percentage points (from 28 percent in 2016 to 34 percent in 2021) for the non-PSNP households.

In Figure 4.4 we bring in the regional dimension to this disaggregation and show patterns of food gap for the PSNP households by survey round. We highlight two takeaways from figure 4.4. First, a large share of PSNP households in Afar (e.g., 45 percent at end-line) and Somali (31 percent at end-line) reported zero food gap and PSNP households in Somali are less likely to report zero food gap than PSNP households in Afar – 10 percentage points lower at baseline and 14 percentage points lower at end-line. Second, between baseline and end-line, the share of PSNP households reporting a food gap of three months and above is also higher for Somali than for Afar; and has declined by 4 percentage points for PSNP households in Afar (from 31 percent in 2016 to 27 percent in 2021) and by 2 percentage points for PSNP households in Somali (from 42 percent in 2016 to 40 percent in 2021).

Next, we look at how female headed households (compared to male headed households) fare in the Lowlands. Figure 4.5 reports mean food gap for end-line by gender of head, region and PSNP status.

Figure 4.5: Mean food gap for end-line by gender of household head, region, and PSNP status



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

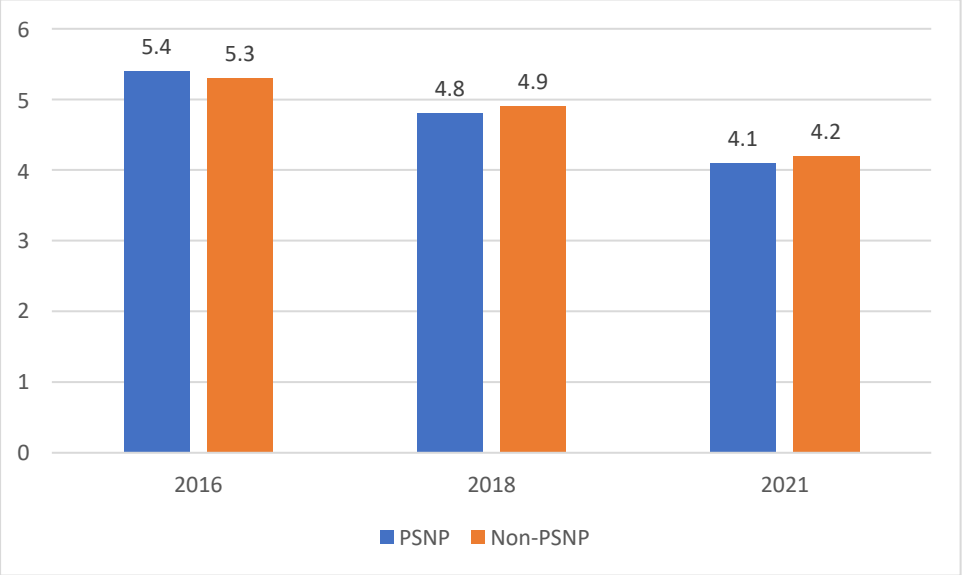
Again, at end-line, regardless of gender of the household head, we note a clear difference between households in Afar and Somali – both male and female headed households in Somali report higher food gap than those in Afar. In addition, female headed non-PSNP households in Afar and Somali tend to report higher food gap than female and male headed PSNP households in their respective regions.

At end-line, the highest mean food gap is reported by non-PSNP male headed households in Somali (Figure 4.5).

4.3 Trends in household dietary diversity

This section reports trends in household diet diversity score (HDDS) constructed from a household consumption module that records the types of food groups from which households consume in a given time period. To construct the index, all food is categorized into 12 food groups: cereals, root and tubers, vegetables, fruits, meat, poultry and offal, eggs, fish and seafood, pulses, legumes and nuts, milk and milk products, oils and fats, sugar and honey, and miscellaneous foods.

Figure 4.6: Mean household dietary diversity score by PSNP status and survey round.

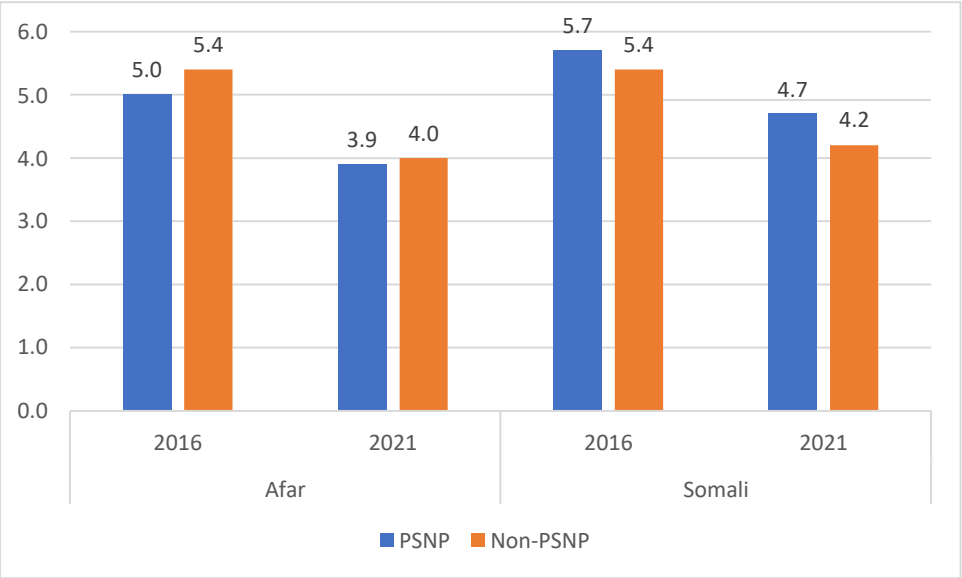


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

A value of one is assigned if the household has consumed from a given food group and zero otherwise, yielding a diet diversity score that ranges between 0 and 12. The higher the index, the better the diet diversity and the vise-versa. This is constructed for baseline and end-line rounds, and the findings are summarized here. Figure 4.6 gives the mean household dietary diversity score by PSNP status and survey round.

At an average dietary diversity score of 4 out of 12 food groups, the diversity of foods consumed in the Lowlands is among the most monotonous.

Figure 4.7: Mean household dietary diversity score by PSNP and non-PSNP households by region and by survey round.



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Even more worrisome is that the mean HDDS declines between baseline and end-line by 1.3 food groups for the PSNP and by 1.1 food groups for the non-PSNP households. A simple mean t-test shows these differences are statistically significant at the 5 percent significance level. We also note that mean household dietary diversity has fallen between baseline and midline. A possible reason for this decline maybe the fact that in the last three years households in the Lowlands have been subjected to several shocks including drought, floods, and pests (see Section 5.4) and that PSNP transfers were either delayed, made irregular, or less than full entitlements (Chapter 6, Performance Report).

Figure 4.6 further breaks down this by PSNP beneficiary status, region, and survey year. Overall, PSNP households from Somali tend to consume slightly more diverse food groups than those from Afar.

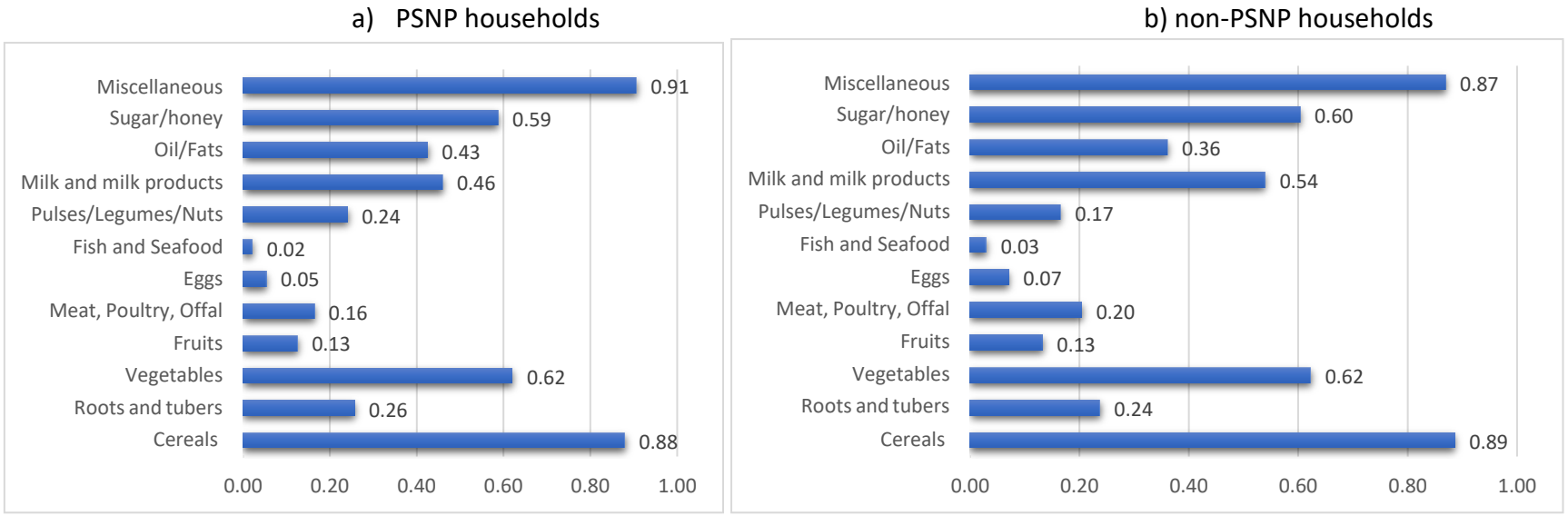
A key observation from these figures is also that, for both PSNP and non-PSNP households in Afar and Somali, mean HDDS has consistently declined between baseline and end-line. This is also consistent with the general declining trend of mean HDDS reported at midline. Specifically, for PSNP households, mean HDDS has declined by 1.1 food groups in Afar and 1.0 food groups in Somali . For the non-PSNP households, mean HDDS has declined by 1.4 food groups in Afar and by 1.2 food groups in Somali (Figure 4.6).

While the mean household dietary score is one of lowest and shows a declining trend, it is also important to identify which of the 12 food groups are major food groups for these households and whether these food groups vary by PSNP status. Figure 4.7 reports percent of households consuming from different food groups in the 7 days preceding the end-line survey day.

Looking at these two figures for PSNP and non-PSNP, we observe a striking similarity and consistency of foods consumed by both the PSNP and non-PSNP households. Consistent with mean HDDS reported

earlier, the vast majority of households (about 90 percent) consume cereals. We also note that almost all households consume from the miscellaneous food groups. Clearly, the four food groups consumed by most households are cereals, miscellaneous, sugar or honey, and vegetables. Among notable in the Lowlands is also that large proportion of PSNP (46 percent) and non-PSNP (54 percent) households consume from milk and milk products, which may as well be seasonal in nature. To the contrary, it is important to note that only a negligible fraction of households consume from animal sourced foods (mainly meat, poultry, offal; eggs; fish and seafoods) and fruits.

Figure 4.8: Percent of households consuming from different food groups at end-line, by PSNP status[4].



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

[4] Miscellaneous food group include spices, herbs, coffee, tea, diet soft drinks, and so on.

4.4 Impact of the PSNP on food security and household dietary diversity

This section summarizes the findings from the impact estimation approach described in chapter 2. The evaluation implemented a cross-sectional difference-in-difference method that uses the nearest neighbor matching estimator. Here, we present summaries of the impact of participation in the PSNP on food security indicated by the food gap and household dietary diversity indicated by the household dietary diversity score. Descriptive trends of these outcomes are reported in sections 4.2 to 4.3. Table 4.1 reports the estimates of the impact of the PSNP on these outcomes. In sum, we do not find impact of the PSNP on food gap and household dietary diversity score. In other words, we cannot reject the null hypothesis that the PSNP did not have statistically significant impact on food security and household dietary diversity outcomes. Finally, we note that the lack of impact on these outcomes is not surprising given PSNP transfers were not timely, unpredictable, and entitlements were incomplete (see chapter 6 of the Performance Report) on the face of shocks reported in chapter 5 (of the Outcomes Report). Moreover, program implementation challenges were further complicated by COVID-19, overall instability, and other broader macroeconomic challenges faced in the last three years.

Table 4.1: Impacts of participation in the PSNP on food security and household dietary diversity, Nearest Neighbor Matching

Outcome	Treatment effect (SE)	Interpretation
Food gap N=3,399	-0.254 (0.241)	We cannot reject the hypothesis that PSNP had no impact on this outcome. The estimated impact is not statistically different from zero.
Household dietary diversity score N=3,283	-0.239 (0.157)	Same as above.

Note: The impact is estimated using a difference-in-difference approach combined with a kernel matching method. Standard errors (SE) are estimated using bootstrap methods with 500 repetitions. Asterisks indicate statistical significance: ***= p<0.01, **= p<0.05, and *= p<0.

Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

4.5 Summary

This chapter has summarized the average trends in food security and household dietary diversity between baseline and end-line. It also gives the key findings of the impact of the PSNP on these outcomes. This section summarizes the main takeaways from this chapter.

- There is substantial heterogeneity in mean food gap between Afar and Somali. Between baseline and end-line, mean food gap has decreased for all households in Afar and PSNP households in Somali, but increased for non-PSNP households in Somali. Specifically, among PSNP beneficiaries, mean food gap has declined by 0.4 month in Afar and 0.6 month in Somali. Mean food gap has also declined by 0.2 month among non-PSNP households in Afar but increased by 1.2 months (an increase by 63 percent) among non-PSNP households in Somali.
- We also note that, overall, mean food gap is higher in Somali than in Afar, regardless of PSNP status.

- Between baseline and end-line, the share of households reporting food gap of three months and above has declined by 2 percentage points for PSNP households but increased by 6 percentage points for the non-PSNP households.
- Female headed non-PSNP households in Afar and Somali tend to report higher food gap than female headed and male headed PSNP households in their respective region.
- For both PSNP and non-PSNP households in Afar and Somali, mean HDDS has consistently declined between baseline and end-line
- PSNP households in Somali consume a slightly more diverse diet (as measured by the number of food groups) than those in Afar. Mean HDDS has declined for PSNP households in both regions.
- Close to 90 percent of households consume cereals and the proportion of households consuming animal sourced foods and fruits is low. No noticeable difference is observed between PSNP and non-PSNP households in terms of the food group types from which they consume.
- Finally, impact estimation results show that participation in the PSNP did not have statistically significant impact on food gap and household dietary diversity scores. It is important to read these results in the context of several unfavorable macroeconomic conditions observed in the last three years (including COVID-19, instability, droughts, floods, and pests) some of which have broader implications to program performance (e.g., poor payment performance) outlinedin the Performance Report chapters and household level shocks (e.g., drought, flood and pests) discussed in Chapter 5 of this Outcomes Report.

Chapter 5: Household Assets, Shocks, and Resilience

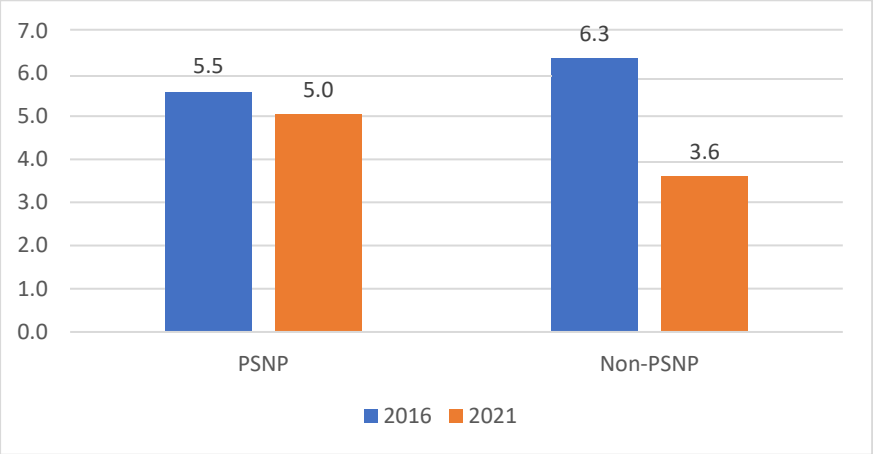
5.1 Introduction

This chapter presents descriptive trends on household assets, shocks, and distress asset sales. It also gives a summary of the empirical findings on the same outcomes. On assets, we focus on livestock, productive assets, and consumer durables. Livestock asset is measured in Tropical Livestock Units (TLU)[5] as well as in values or monetary units. Other assets are also presented in monetary values. Asset values are adjusted for inflation to make them comparable across periods. We specifically compare baseline (or 2016) mean outcomes against end-line (2021) mean outcomes and present changes in these outcomes along with t-tests as required. We disaggregate our results by region, PSNP status and gender of the household head. In the empirical analysis, a matching estimator that takes into account the cross-sectional time series nature of the lowland data is implemented (see chapter 2 for methodological details). We begin with descriptive trends and present the empirical results at the end.

5.2 Trends in livestock asset holdings

This section summarizes the mean changes in livestock assets holdings of PSNP and non-PSNP households between the baseline and end-line. We further disaggregate these results by gender of the household head and region. We first focus on livestock assets using the TLU measure and Figure 5.1 reports the trends in TLU by PSNP and survey round.

Figure 5.1: Mean TLU by PSNP and survey round



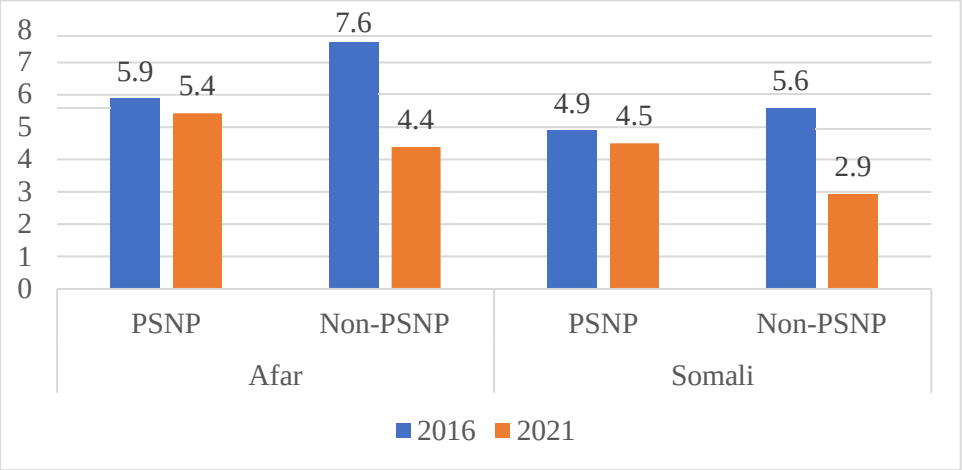
Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

[5] A TLU is a convenient way of adding different livestock types into a single measure. The standard measure of a TLU is an animal with live weight of 250 kg (Jahnke (1982)). TLU are expressed as ratios to this, the ratios being based on metabolic weights. So, for example, six sheep have the same energy requirements as one head of cattle and so six sheep are one TLU (or, put another way, 1 sheep = 0.15 TLU).

We see that between baseline and end-line, mean TLU has decline for all Lowland households but more dramatically so for the non-PSNP households in the Lowlands – a decline by 43 percent for the non-PSNP but only by 9 percent for the PSNP households. How do households in Afar and Somali fare in terms of this decline in TLU?

Figure 5.2 further disaggregates the mean TLU by region, survey round and PSNP status. Disaggregation unmasks striking heterogeneities across regions and PSNP status. Clearly, there is a dramatic decline in mean TLU for the non-PSNP households in Afar (by 42 percent) and Somali (by 48 percent)- mean TLU has declined by 6 percentages points more for non-PSNP households in Somali than those in Afar. Mean TLU for the PSNP households in both regions has declined only by 8 percent (Figure 5.2).

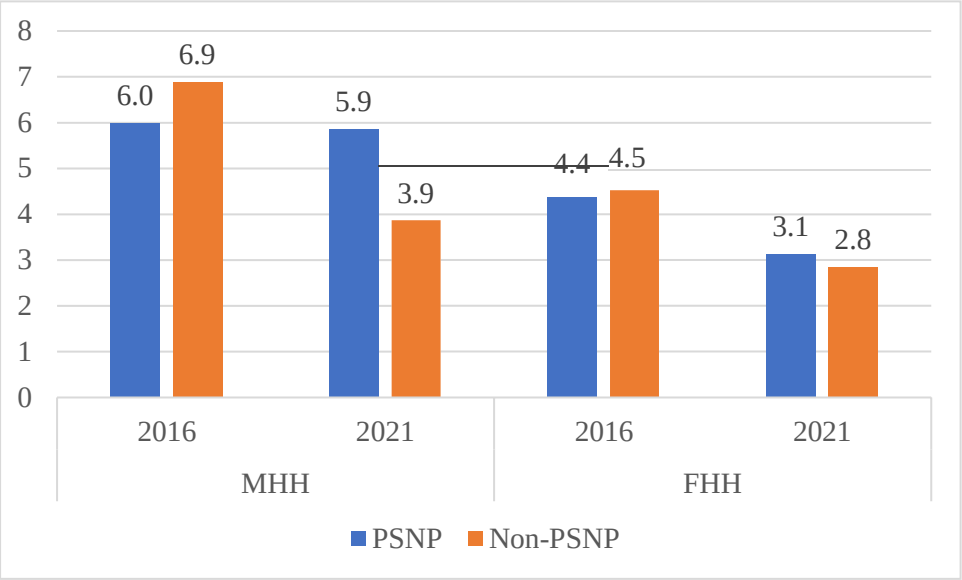
Figure 5.2: Mean TLU by PSNP, survey round, and region



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

To what extent are these trends and heterogeneities in TLU common among female versus male headed households? Figure 5.3 breaks down the mean TLU by gender of the household head, survey round and PSNP status. A striking result to note when comparing male versus female headed households between rounds, female headed households owned consistently lower mean TLUs than male headed households. Overall, at baseline, female headed households owned 45 percent lower mean TLUs than male headed PSNP households. This gap has further deteriorated in 2021 where, on average, female headed PSNP households owned 60 percent less TLU of that of male headed household. Compared to male headed non-PSNP, female headed non-PSNP households have also owned consistently lower TLUs on average (36 percent lower in 2016 and 52 percent lower in 2021) (Figure 5.4). A drawback in using TLU is that it does not show the exact value of the livestock holding in monetary terms.

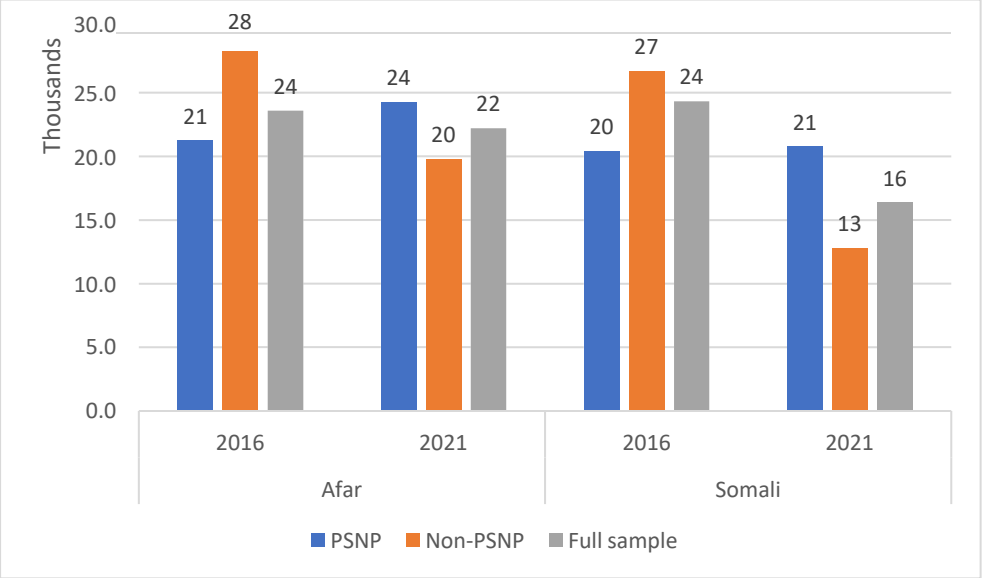
Figure 5.3: Mean TLU by PSNP status, survey round and gender of head



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Next, we present the total value of livestock holding at the household. Figure 5.4 shows the mean total value of livestock assets owned (in thousands of Birr[6]) by PSNP status, region, and survey round. Consistent with the mean TLU trends, mean value of livestock has slightly increased for PSNP households in Afar but remained roughly the same in Somali.

Figure 5.4: Mean total value of livestock assets owned, by PSNP status, survey year and round (Thousands, Birr).



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

[6] $US 1 was equivalent to 44 Birr at the time of the end-line survey.

Likewise, mean value of livestock for non-PSNP households in Afar and Somali has shown dramatic decline between baseline and end-line (a decline by 43 percent from 28 thousand Birr in 2016 to 20 thousand Birr in 2021 for Afar and a decline by 52 percent from 27 thousand Birr in 2016 to 13 thousand Birr in 2021 for Somali) (Figure 5.4).

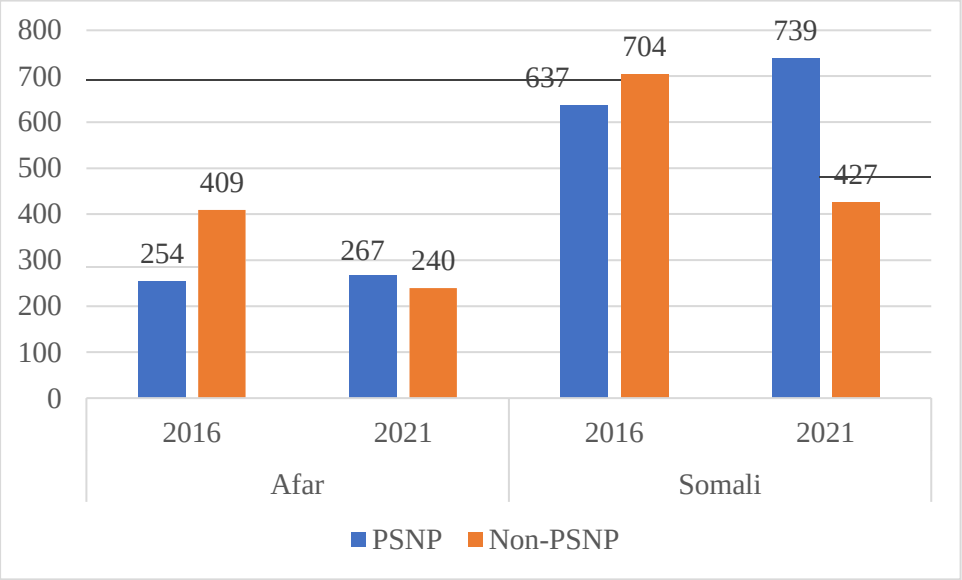
5.3 Trends in values of productive assets, consumer durables, and total assets

As indicated in the introduction, this section presents trends in mean value of productive assets, consumer durables and total assets. Our surveys included elaborated modules asking respondents whether and how many of these household assets they owned at the time of the survey. They were then asked to give values of these assets if they were to sell them at market rates. This information is thus used to construct the aggregated values of each of these assets at the household level and are subsequently employed in analyses after adjusting them for inflation.

5.3.1 Trends in productive assets

We begin with Figure 5.5 that reports the mean value of productive assets by PSNP status, survey round and region. Note again that these values are adjusted for inflation using a consumer price index for the lowlands. We see that the mean value of productive assets for both rounds and all household types are lower in Afar than in Somali region. Specifically, at end-line, the mean value of productive assets for all households in Afar was a little more than half of that of Somali. On the other hand, across the two rounds, the mean value productive assets owned by PSNP households in Afar has remained roughly the same whereas for PSNP households in Somali this has increased from 637 Birr to 739 Birr.

Figure 5.5: Mean value of productive assets (Birr), by PSNP status, round and region



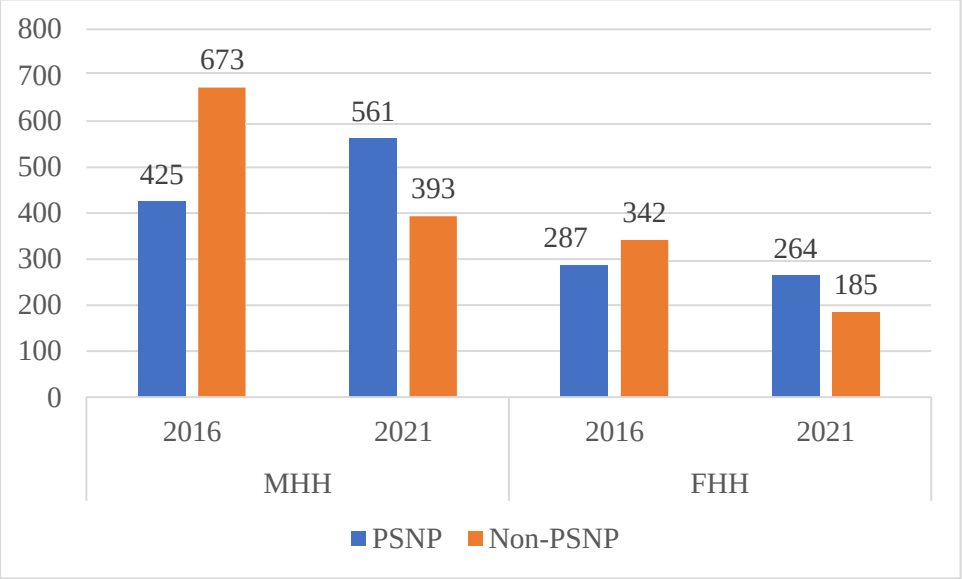
Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

We also see noticeable differences among household types within regions and rounds. For instance, the mean value of productive assets for non-PSNP households in both regions has fallen substantially (by 70 percent in Afar and 65 percent in Somali) (Figure 5.5).

5.3.2 Female- versus male-headed households

Access to productive assets is critical for production and improving food security and livelihoods. In many cases, not having access or not owning these inputs may predetermine many of the outcomes of interest for the PSNP. This maybe one area in which women headed households find themselves less favored. To understand this issue, we disaggregate the data by gender of household head for the full sample. Figure 5.6 reports a summary of the main results. As evidenced earlier for livestock asset ownership, regardless of their PSNP status, female headed households in the lowlands own marginally lower mean value of productive assets than their male counterparts. For example, at end-line, the average women headed household owned only half of the value of productive assets owned by a male headed household. While the mean value of productive assets of women headed PSNP households remained relatively stable between baseline and end-line, for the non-PSNP women headed household this figure has further deteriorated from 342 Birr at baseline to 185 Birr at end-line.

Figure 5.6: Mean value of productive assets, by PSNP, survey round and gender of head



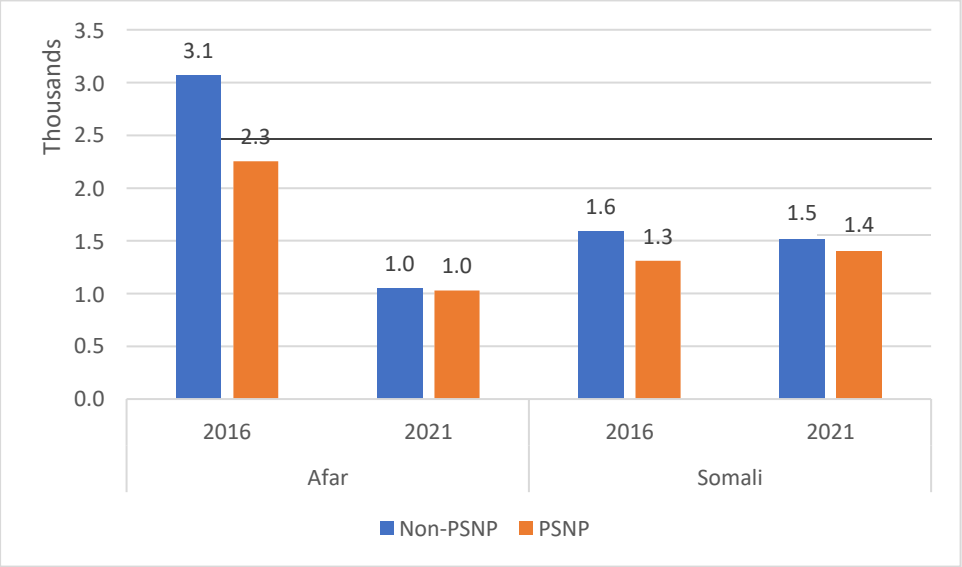
Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

5.3.3 Trends in consumer durables

Improvements in household consumer durables are often signs of improvements in overall wellbeing of the household. This section discusses the key trends in the mean value of household durable assets. Figure 5.7 reports summary of mean value of total consumer

durables (in thousands, Birr) by PSNP status, survey round, and region. Notice that the mean values of these assets for Afar 2016 appear to be a bit peculiarly higher than the rest. In that context, we see a dramatic drop in mean value of consumer durables between baseline and end-line. Midline and end-line trends are consistent (we are unsure about why this drop might have occurred at midline). A similar drop was documented at midline. For Somali, mean values of consumer durables remained stable across rounds as well as the different householdgroups and mean values of non-PSNP households (average of 1350 Birr in both rounds) are slightly lower than the mean values of PSNP households (average of 1550 Birr in both rounds).

Figure 5.7: Mean value of consumer durables, by PSNP status, survey round, and region (Birr)



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

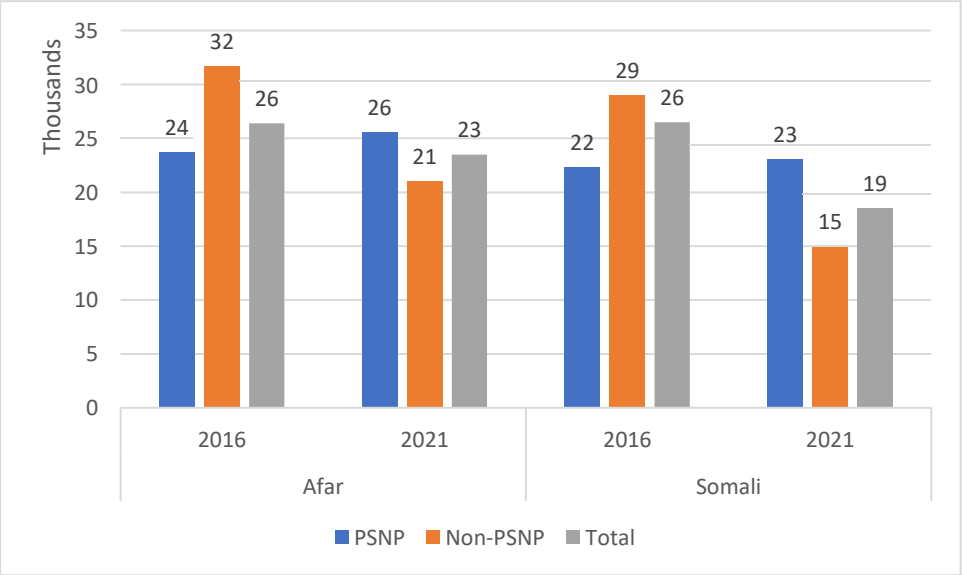
5.3.4 Trends in total household assets

This section brings together all the total assets discussed so far and provides the overall picture of asset holdings across the different household groups, region and rounds. Figure 5.8 shows the mean total value of assets households owned by survey round, PSNP status and region.

There are two key takeaways from this figure. First, between baseline and end-line, mean total value of assets for all households has declined in both Afar (by 12 percent) and Somali (by 27 percent). This is driven by a substantial decline in mean total value of assets for the non-PSNP households in Afar (by 48 percent) and Somali (by 34 percent). Notice that both regions started from similar mean value of total assets (26 thousand) at baseline. Second, in contrast, between baseline and end-line, mean value of total assets for PSNP households has increased by 5 percent in Afar and 8 percent in Somali. Overall, at end-line, an average

household in Somali, regardless of PSNP status, owns lower mean total value of assets than an average household in Afar.

Figure 5.8: Mean total value of household assets, by PSNP status, survey round, and region (Thousands, Birr)



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

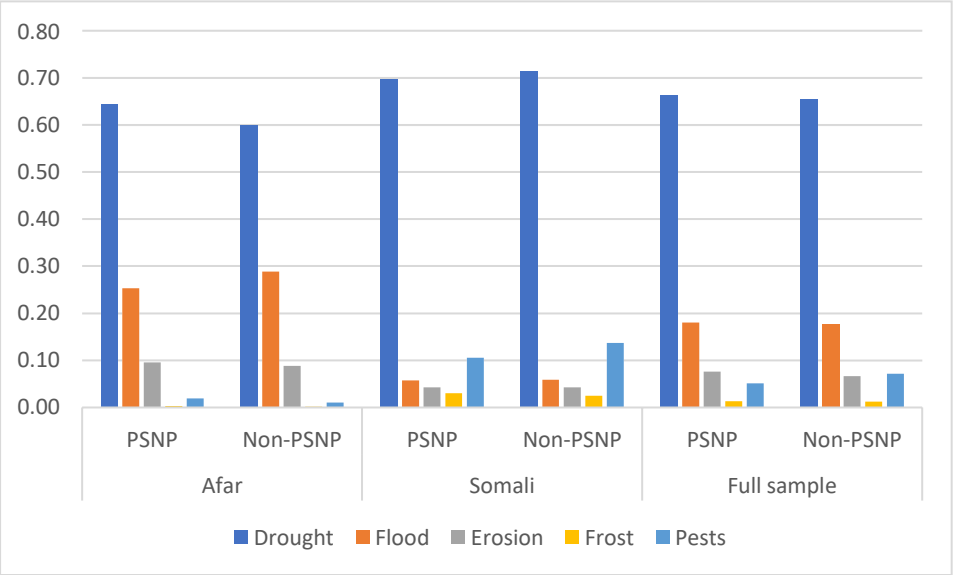
5.4 Shocks and resilience

At the heart of the design of the PSNP4 is a key goal of enhancing households' resilience to shocks by preventing asset depletion through distress sales when shocks hit (FDRE 2014). An important question in this evaluation is therefore whether and to what extent the PSNP4 has prevented distress sales of assets. While this question is addressed in the next section (see section 5.5), in this section, we provide descriptive results of shocks experienced, and trends in asset depletion between baseline and end-line at different levels of disaggregation.

5.4.1 Shocks

To understand whether and what type of shocks households experience over a three-year horizon, we asked our respondent the following question at the end-line: "Thinking about the last three years what were the most important shocks you experienced?"

Figure 5.9: Percent of households that experienced shocks in 2021, by region and PSNP status



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figure 5.9 summarizes the proportion of households reporting each of these five shock types by PSNP status and region. Clearly, of all shocks, drought or lack of moisture stands out as the most important shock for all households in the lowlands, regardless of their PSNP status. Between 60 to 70 percent of all households have reported drought as most important. Households from the Somali sample are slightly more likely to report drought as key shock. Flood or too much rain stands as the second most important shock experienced by all households, but more pronounced in Afar than in Somali (possibly because of the Awash River flooding). A non-negligible proportion of households have reported pest as most important shock. This is not surprising given the locust swarms manifested in many areas of the lowlands in recent years. Finally, it is important to note that there is not much difference between PSNP and non-PSNP households with regards to these shocks.

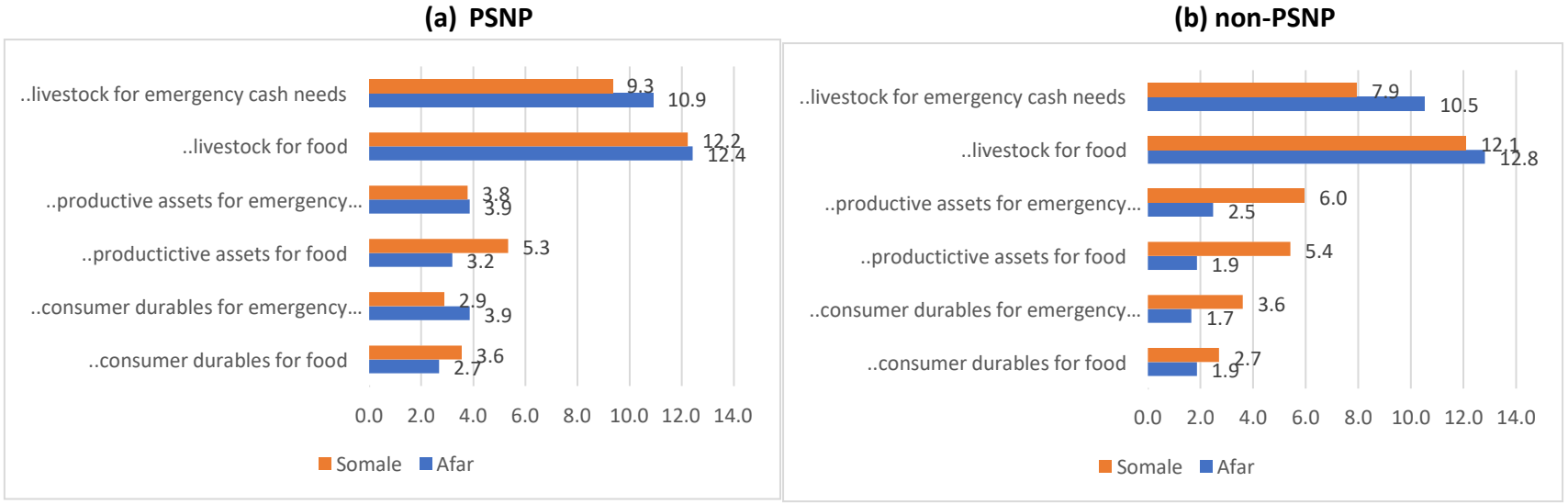
5.4.2 Distress asset sales and resilience

How resilient are PSNP households to shocks and do they sell their productive assets to cope with it? Does resilience vary by PSNP and non-PSNP or region where the household comes from? Figure 5.10 (a) and (b) show the percent of households that reported they were forced to undertake distress sales of any livestock, productive, and consumer durable assets in the last two years by PSNP status and region. These are summarized by two important reasons of distress sales mainly to fulfill food needs (bridge food gap) or satisfy an emergency cash need.

Overall, there are three important points to note. First, there is not much difference between PSNP and non-PSNP households with regards to the proportion of households reporting these distress sales. Second, our samples from Afar and Somali are mirrors of each with regards to

these distress sales. Third, among the three asset types, distress sales of livestock assets for food and emergency cash needs stands out as the highest reported by relatively higher proportion of households (about 8 – 12 percent of all households reporting it).

Figure 5.10 (a): Distress sale of assets by PSNP households at end-line (%), by region



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

5.5 Impact of the PSNP on household assets and resilience

This section summarizes the findings from the impact estimation exercises described in chapter 2. The evaluation implemented a cross-sectional difference-in-difference method that uses the nearest neighbor matching estimator. Here, we present summaries of the impact estimates of participation in the PSNP on household assets, mainly livestock measured using TLU and real values, real value of durable assets, real value of productive assets, and real value of total assets. Descriptive trends of these outcomes are reported in sections 5.2 to 5.4. Table 5.1 gives the estimates of the impact of the PSNP on these outcomes. In sum, we do not find any impact of the PSNP on any of these outcomes at the 95 percent significancelevel. In other words, we cannot reject the null hypothesis that the PSNP had no statistically significant impact on these outcomes. These results do not come as surprises given 60 -70 percent of households reported they had experienced shocks, including drought, flood, and pests. These are over and above to the overall ongoing instability in the country. We also know from the Performance Report that PSNP payments were not timely, not predictable, and not given in full entitlements. In line with this, section 5.4 of this report indicates that 8 - 12 percent of PSNP households have had distress sales of their livestock and productive assets for food and emergency cash needs.

Table 5.1: Impacts of participation in the PSNP on household assets, Nearest Neighbor Matching

Outcome	Treatment effect (SE)	Interpretation
Tropical livestock units (TLU) N=3,399	0.042 (0.093)	We cannot reject the hypothesis that PSNP had no impact on this outcome. The estimated impact is not statistically different from zero.
Real value of livestock holdings N=3,399	0.177 (0.285)	Same as above.
Real value of durable assets N=3,399	0.104 (0.188)	Same as above.
Real value of all assets combined N=3,399	0.098 (0.126)	Same as above.

Note: The impact is estimated using a difference-in-difference approach combined with a kernel matching method. p-values are estimated using bootstrap methods with 500 repetitions. Asterisks indicate statistical significance: ***= p<0.01, **= p<0.05, and *= p<0.1.

Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

5.6 Summary

This chapter presented descriptive trends in household asset holdings, shocks experienced, and distress sales of any of the household assets for food or emergency cash needs. Household assets included livestock (measured by TLU and value in Birr), productive assets, consumer durables, and total assets. It also reported on the impact of PSNP participation on some of these assets. In this section, we provide the key takeaways in line with the research questions asked in the TOR for this work.

- Mean TLU has declined between baseline and end-line for all households but more dramatically so for the non-PSNP households – a decline by 43 percent for the non-PSNP but only by 9 percent for the PSNP households. Regional disaggregation also shows a dramatic decline in mean TLU for the non-PSNP households in Afar (by 42 percent) and Somali (by 48 percent) but for PSNP households in both regions, the decline was only by 8 percent. Mean value of livestock assets followed the same trends.
- Overall, mean value of productive assets for both rounds and all household types are lower in Afar than in Somali.
- At end-line, the average household from Afar owns higher total value of assets (28,000 Birr on average) than the same household from the Somali sample (18,000 Birr on average).
- Female headed households owned consistently lower mean TLUs than male headed households. On average, at baseline, female headed households owned 45 percent fewer TLUs than male headed PSNP households. This gap has widened further in 2021 where, on average, female headed PSNP households owned 60 percent less TLU than that of male headed household.
- Regardless of their PSNP status, the average value of productive assets owned by female headed households in the lowlands is lower than their male headed counterparts.
- Drought or lack of moisture stands out to be the most important shock for all households in the lowlands, regardless of their PSNP status. Between 60 to 70 percent of all households have reported drought as the most important.
- There is little difference between PSNP and non-PSNP households with regards to the proportion of households reporting distress sales and, among the three asset types, distress sales of livestock for food and emergency cash needs stands out as the highest reported (about 8 – 12 percent of all households reporting it).
- Finally, impact estimation results show that the PSNP had no statistically significant impact on any of the household assets discussed in this chapter, mainly livestock TLU, real value of livestock, real value of durable assets, and real value of total assets.

Chapter 6: Labor Allocation and Work Intensity

6.1 Introduction

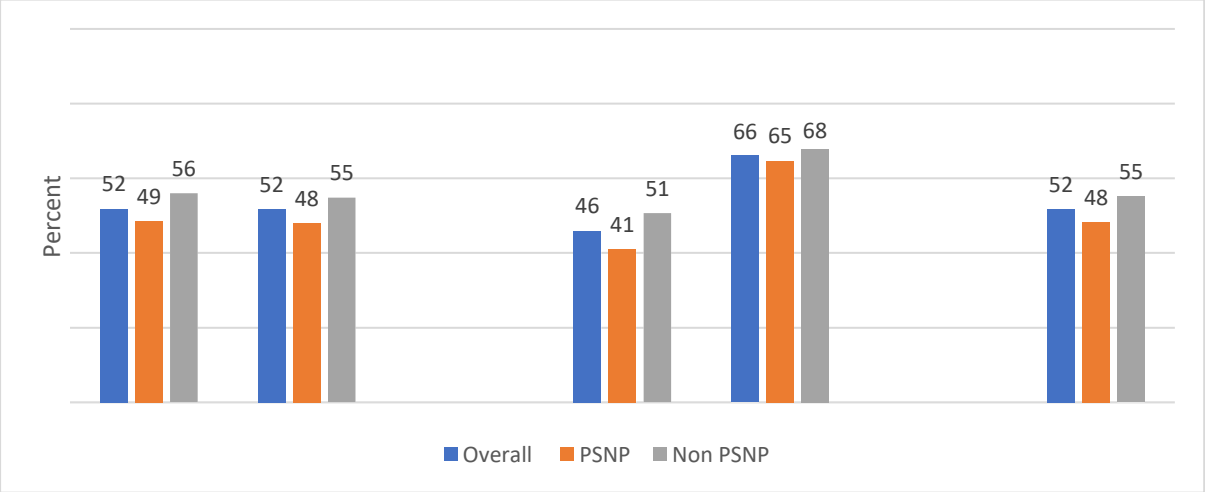
In this chapter, we present descriptive results on labor allocation and work intensity at end-line.

This chapter tries to address **RQ11: Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?** However, since these data were only collected at end-line and because of the impact estimation strategy in the lowlands we cannot measure impacts of the PSNP on these outcomes. Therefore, we can only partially address this question for the lowlands.

6.2 Labor allocation and work intensity

The end-line survey collected information on labor allocation in the past 7 days by all household members. Overall, 46 percent of the households reported not having spent any time on any activities – related to agricultural or non-agricultural work, casual labor, salaried work or the public works – in the 7 days prior to the survey. We find that PSNP households are less likely to report not having worked in the past 7 days as compared to non-PSNP households. We do not observe a lot of regional differences, as seen in Figure 6.1. About 52% households in Afar and 52% households in Somali reported not having worked in the last 7 days. The difference in having worked or not in the previous 7 days is more pronounced when we compare male-headed and female-headed households (Figure 6.1), with 46 percent of male-headed households reporting they spent zero hours on these activities in the past 7 days as compared to 66% percent among the female-headed households.

Figure 6.1: Households that did not work in the past 7 days, by region, headship and PSNP status

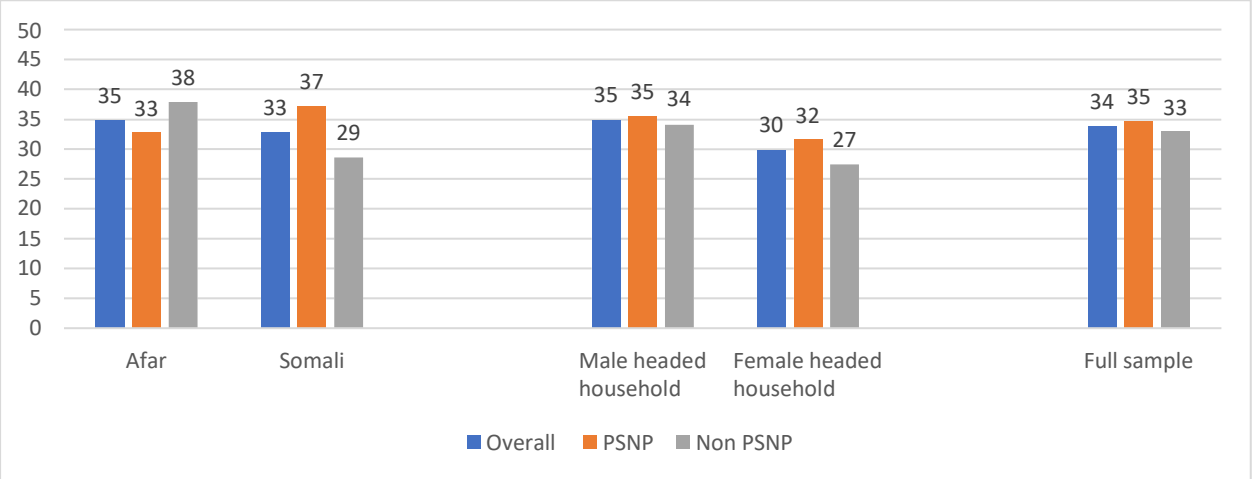


Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Conditional on spending any time on labor activities, the average total numbers of hours spent by the household in the 7 days prior to the survey is 34 hours. There is some variation in total hours spent on labor activities across regions with Afar at 35 hours and Somali at 33 hours (Figure 6.2). Fewer hours are spent by PSNP households as compared to non-PSNP households . Whereas in Somali, we observe

a big difference in the opposite direction– PSNP households spent an average of 37 hours on various activities in the last 7 days whereas non-PSNP households spend 29 hours on average. We also observe that male-headed households spend 35 hours on average in the previous 7 days as compared to 30 hours among the female-headed households.

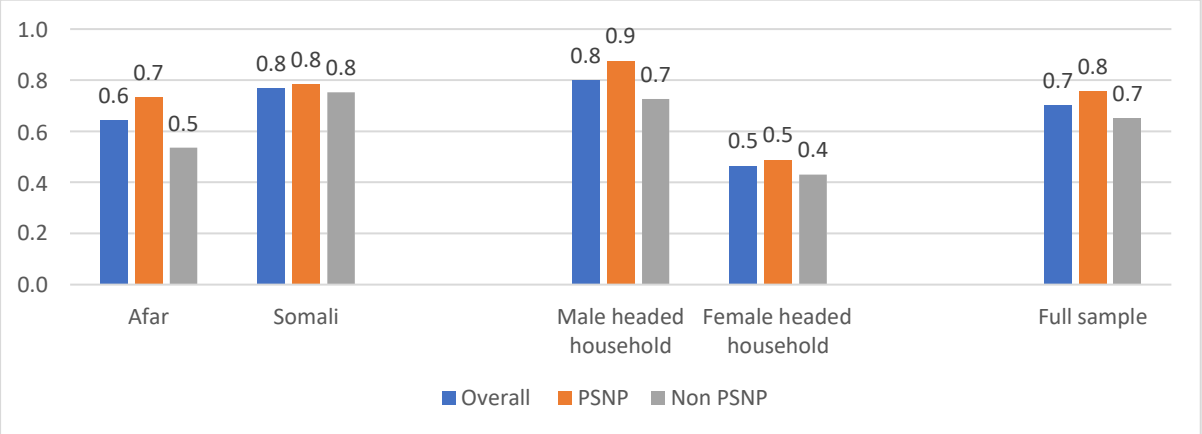
Figure 6.2: Total hours spent by households on labor activities, by region, headship and PSNP status



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

The households were asked about their labor allocation in the 7 days prior to the survey in five types of activities. These included agriculture (including livestock and fishing-related activities), non-agricultural business, casual/part-time labor, wage or salaried labor, and; work on PSNP public works. On average households were engaged in 0.7 out of these 5 activities in the last 7 days. This is slightly higher at 0.8 among the PSNP households as compared to 0.7 among non-PSNP households with some variation across regions and sex of the household head (Figure 6.3).

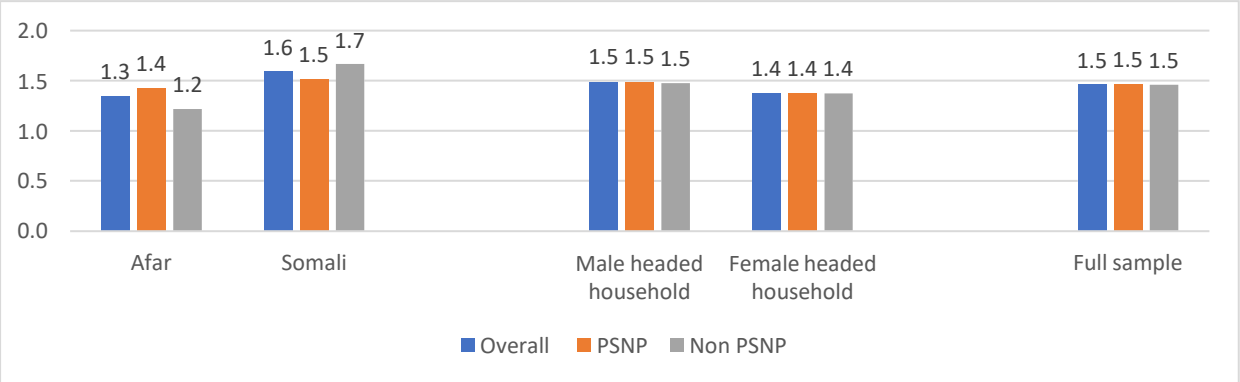
Figure 6.3: Total number of activities involved in last 7 days, by region, headship and PNSP status



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

When we restrict the sample to those who were engaged in any activities the last 7 days, these averages increase to 1.5 activities overall and we do not observe a difference between PSNP households and non-PSNP households (Figure 6.4). And the differences by region and sex of household head are also less pronounced.

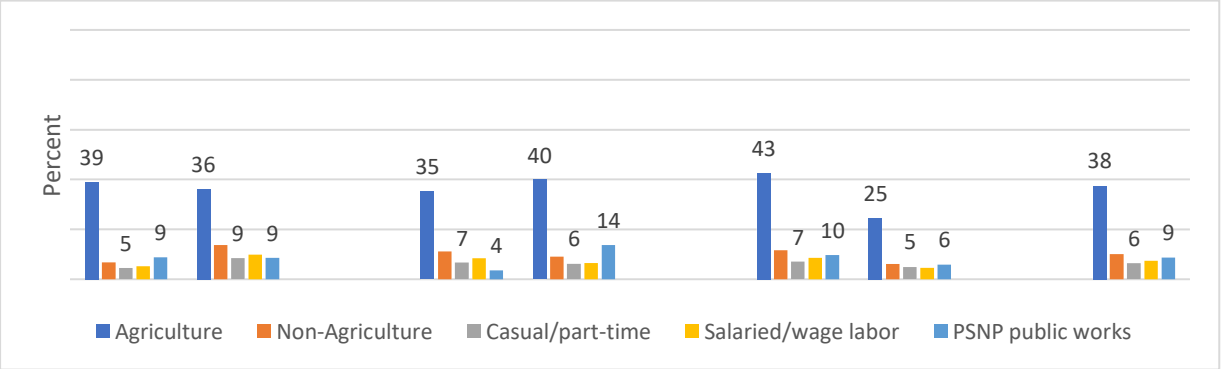
Figure 6.4: Among those involved in any activity total number of activities involved in last 7 days, by region, headship and PNSP status



Source: Authors' computation using PSNP4 Baseline End-line (2021) Household Survey.

On average 38% of the households reported being engaged in agricultural activities in the last 7 days (Figure 6.5). The rest of the activities are reported by a small percentage of households. PSNP public works is reported by 9% of the households and as expected this is concentrated among the PSNP households.

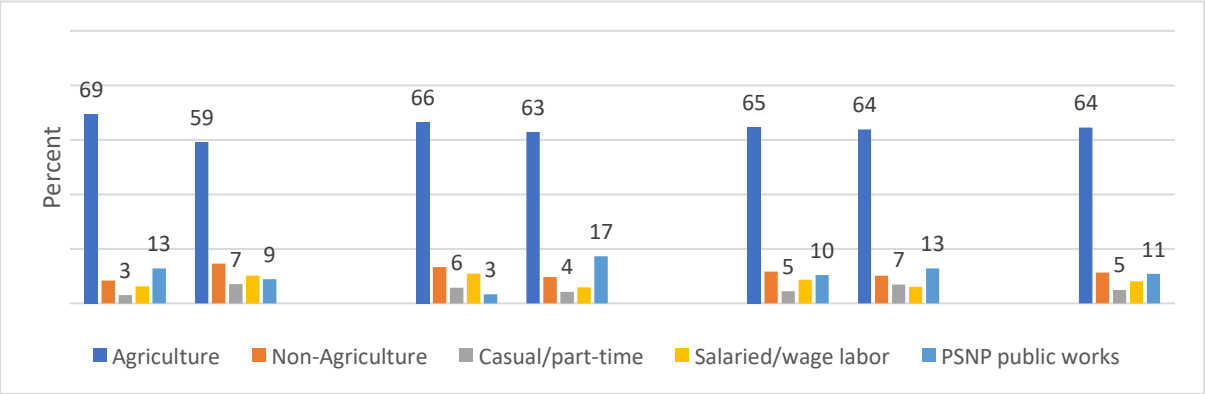
Figure 6.5: Percentage of households engaged in different activities, by region and headship



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

We also construct an indicator that shows the relative level of effort in the 5 different activities in the past 7 days. This is the percentage of total hours spent during the last 7 days that were allocated to a specific activity. This is shown in Figure 6.6 unsurprisingly revealing that agricultural activities form the bulk of the time these households expended in the previous 7 days.

Figure 6.6: Average percentage of time spent on different activities by households, by region and headship



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Table 6.1 presents summary statistics for the labor outcome variables by PSNP status and also presents the p-value on the difference in the means across the two groups. We find that overall PSNPhouseholds are more likely to have worked in the 7 days prior to the survey, they are also engaged ina higher number of activities (some of this is by definition since 1 out of the 5 activities is available only to this group) and spend more hours.

Table 6.1: Labor allocation and intensity outcomes, by PSNP status in 2021

	PSNP households			Non-PSNP households			p-value
	Mean	SD	N	Mean	SD	N	
Total hours spent by household, 7 days	17.90	27.74	1,045	14.67	26.78	1,033	0.007
No. activities HH engaged, 7 days	0.76	0.97	1,045	0.65	0.97	1,033	0.013
HH engaged in ag activities in last 7 days	0.40	0.49	1,045	0.35	0.48	1,033	0.022
HH engaged in non-ag activities in last 7 days	0.09	0.29	1,045	0.11	0.32	1,033	0.107
HH engaged in casual labor in last 7 days	0.06	0.24	1,045	0.07	0.25	1,033	0.670
HH engaged in salaried work in last 7 days	0.07	0.25	1,045	0.08	0.28	1,033	0.116
HH engaged in PW in last 7 days	0.14	0.34	1,045	0.04	0.19	1,033	0.000
Percent hours HH engaged in ag activities last 7 days	0.63	0.43	540	0.66	0.43	461	0.195
Percent hours HH engaged in non-ag activities last 7 days	0.10	0.25	540	0.13	0.28	461	0.026
Percent hours HH engaged in casual labor last 7 days	0.04	0.16	540	0.06	0.18	461	0.153
Percent hours HH engaged in salaried work last 7 days	0.06	0.20	540	0.11	0.27	461	0.001
Percent hours HH engaged in PW last 7 days	0.17	0.34	540	0.03	0.16	461	0.000

Notes: SD refers to standard deviations. p-values are from a t-test for of equality of means of the outcome across the PSNP and non-PSNP samples. The normalized difference for each outcome is defined as the ratio of the PSNP sample mean minus the non-PSNP sample mean to the square root of one half of the sum of the PSNP sample and non-PSNP sample variances.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Table 6.2: Labor allocation and intensity outcomes, by region and gender of household head in 2021

	Afar		Somali		Difference by region (p-value)	Female-headed		Male-headed		Difference by gender of HHH (p-value)
	Mean	N	Mean	N		Mean	N	Mean	N	
No. of hours HH engaged in any activity, 7 days	16.745	1,079	15.800	999	0.431	10.026	587	18.757	1,491	0.000
No. activities HH engaged, 7 days	0.645	1,079	0.768	999	0.004	0.462	587	0.799	1,491	0.000
HH engaged in ag activities in last 7 days	0.390	1,079	0.360	999	0.161	0.245	587	0.427	1,491	0.000
HH engaged in non-ag activities in last 7 days	0.069	1,079	0.137	999	0.000	0.061	587	0.117	1,491	0.000
HH engaged in casual labor in last 7 days	0.045	1,079	0.085	999	0.000	0.049	587	0.070	1,491	0.079
HH engaged in salaried work in last 7 days	0.053	1,079	0.099	999	0.000	0.046	587	0.087	1,491	0.002
HH engaged in PW in last 7 days	0.088	1,079	0.086	999	0.874	0.060	587	0.098	1,491	0.005
Percent hours HH engaged in ag activities last 7 days	0.694	519	0.592	482	0.000	0.638	197	0.647	804	0.815
Percent hours HH engaged in non-ag activities last 7 days	0.084	519	0.146	482	0.000	0.101	197	0.117	804	0.452
Percent hours HH engaged in casual labor last 7 days	0.031	519	0.071	482	0.000	0.070	197	0.046	804	0.070
Percent hours HH engaged in salaried work last 7 days	0.063	519	0.102	482	0.009	0.062	197	0.087	804	0.189
Percent hours HH engaged in PW last 7 days	0.128	519	0.088	482	0.026	0.128	197	0.104	804	0.281

Notes: SD refers to standard deviations. HHH = Household Head. p-values are from a t-test for of equality of means of the outcome across the PSNP and non-PSNP samples. The normalized difference for each outcome is defined as the ratio of the PSNP sample mean minus the non-PSNP sample mean to the square root of one half of the sum of the PSNP sample and non-PSNP sample variances.

Source: Authors' computation using PSNP4 End-line (2021) Household Survey

6.3 Child labor

In this subsection we present summary statistics on work performed by children 7-14 years old and those who are 15-17 years old in the 7 days prior to the survey. The work activities include agricultural activities, non-agricultural activities, casual/part-time work, salaried and PSNP public works.

We present the descriptive statistics for children 7-14 years old in Table 6.3. We observe that about 20% PSNP households reported their children being engaged in agricultural activities in the last 7 days as opposed to 18% reporting the same among non-PSNP households (difference is not statistically significant). A very small percentage of households, ranging from 1-3%, report children being involved in other work activities. Additionally, PSNP households are less likely to report their children being involved in casual labor and salaried work as compared to non-PSNP households.

Table 6.3: Child labor among 7-14 years old in the 7 days prior to survey, end-line

	PSNP		Non-PSNP		p-value of difference by PNSP status
	Mean	N	Mean	N	
Engaged in agricultural activities	20%	705	18%	575	0.285
Total hours spent in ag activities in last 7 days	0.84	160	0.78	119	0.189
Engaged in non-ag activities	3%	705	3%	575	0.425
Total hours in non-ag activities in last 7 day	0.08	160	0.09	119	0.829
Engaged in casual labor	1%	705	3%	575	0.010
Total hours in casual labor in last 7 day	0.03	160	0.08	119	0.027
Engaged in salaried work	0%	705	2%	575	0.020
Total hours in salaried work in last 7 day	0.01	160	0.04	119	0.017
Engaged in PW	1%	705	1%	575	0.967
Total hours in PW in last 7 days	0.05	160	0.02	119	0.150

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Next, we present the descriptive statistics for children 15-17 years old in Table 6.4. We observe that about 24% of PSNP households reported their children being engaged in agricultural activities in the last 7 days as opposed to 16% among non-PSNP households who reported the same (difference is statistically significant).

It is interesting to note that, according to respondents, the total amount of time children spent working in the 7 days covered by the survey is less than an hour in all cases.

Table 6.4: Child labor among 15-17 years old in the 7 days prior to survey, end-line

	PSNP		Non-PSNP		Difference by PNSP status (p-value)
	Mean	N	Mean	N	
Engaged in agricultural activities	24%	387	16%	249	0.023
Total hours spent in ag activities in last 7 days	0.79	107	0.75	48	0.582
Engaged in non-ag activities	2%	387	4%	249	0.222
Total hours in non-ag activities in last 7 day	0.04	107	0.15	48	0.004
Engaged in casual labor	2%	387	2%	249	0.959
Total hours in casual labor in last 7 day	0.03	107	0.03	48	0.801
Engaged in salaried work	2%	387	2%	249	0.666
Total hours in salaried work in last 7 day	0.02	107	0.06	48	0.078
Engaged in PW	5%	387	1%	249	0.018
Total hours in PW in last 7 days, F	0.13	107	0.01	48	0.013

Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

6.4 Summary

In this chapter we presented descriptive results on labor allocation and work intensity in the 7 days prior to the end-line survey. The activities included in this section are agriculture (including livestock and fishing-related activities), non-agricultural business, casual/part-time labor, wage or salaried labor, and work on PSNP public works.

- 46 percent of the households reported not having spent any time on any activities – related to agricultural or non-agricultural work, casual labor, salaried work or public works – in the 7 days prior to the survey. We observe differences by region, by sex of household head and PNSP status.
- PSNP households are less likely than non-PSNP households to report not having worked in the past 7 days.
- On average households were engaged in less than 1 out of 5 activities in the last 7 days.
- On average 43% of the households reported being engaged in agricultural activities in the last 7 days. The rest of the activities are reported by a small percentage of households.
- PSNP public works participation by children is reported by 6% of PSNP households.
- PSNP households spent a larger number of hours on various activities in the last 7 days as compared to non-PSNP households.

Given the data constraints we cannot answer the research question, **RQ11: Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?** But we do find that PSNP households are more likely to be working and work more hours as compared to non-PSNP households. This may be reflective of their relative economic status which may require the PSNP households to work more. However, this is suggestive evidence that PSNP is not leading to a reduction in work effort.

Chapter 7: Poverty

7.1 Introduction

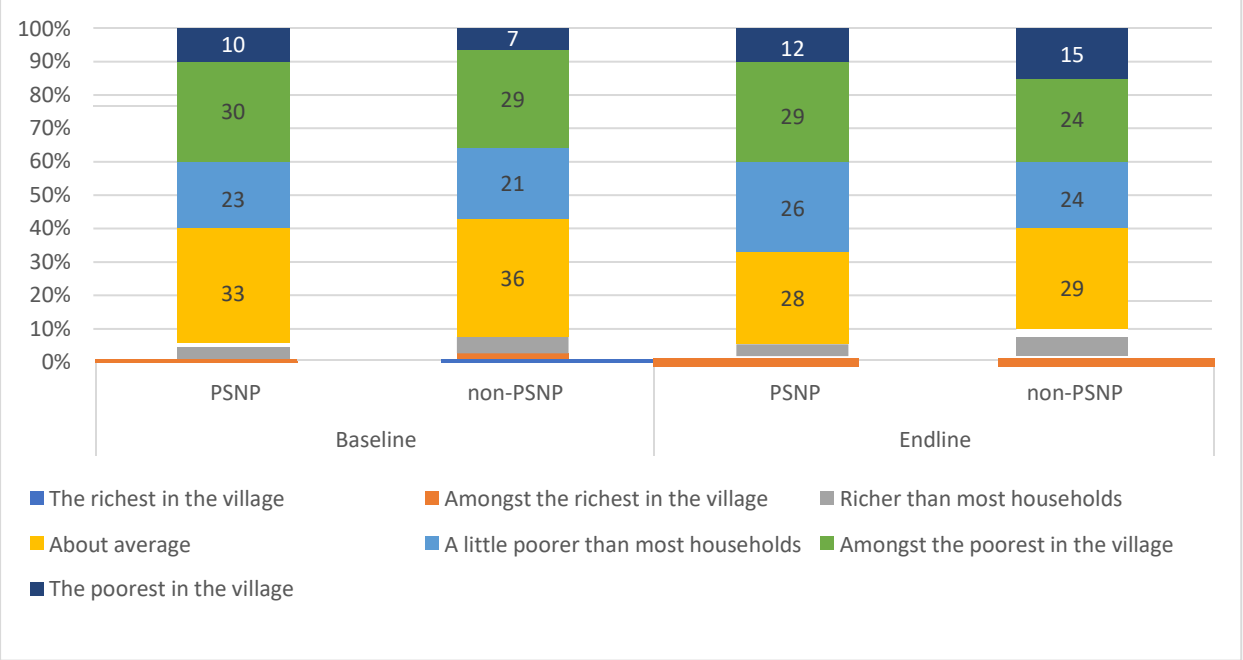
In this chapter, we present results on poverty using subjective measures. We start with descriptive results showing trends over time and by region and sex of the household head. We then present impact of the participating in the PSNP on subjective poverty. This chapter will address **RQ3: Has PSNP reduced poverty?**

In addition to poverty measures, we also provide descriptive results on indicators for agency.

7.2 Subjective wellbeing

In the survey, data was collected on the self-perceived wellbeing among households. Respondents were asked how they would describe their household as compared to other households in the village. The option included – the richest, among the richest, richer than most, about average, a little poorer than most, amongst the poorest, the poorest. As can be seen in Figure 7.1, PSNP household's perception of their economic standing is lower than the non-PSNP households at baseline and at end-line. We observe some movement in this perception over time. At baseline, 33% of PSNP households and 36% of non-PSNP households reported themselves as "about average" relative to others in the village. This percentage has gone down in both groups and the percentage reporting "little poorer than most" and "the poorest" has gone up. Among the non-PSNP households, the percentage reporting themselves as "the poorest" has more than doubled during this time.

Figure 7.1: Self-perceived economic standing in village, by PSNP status and survey round

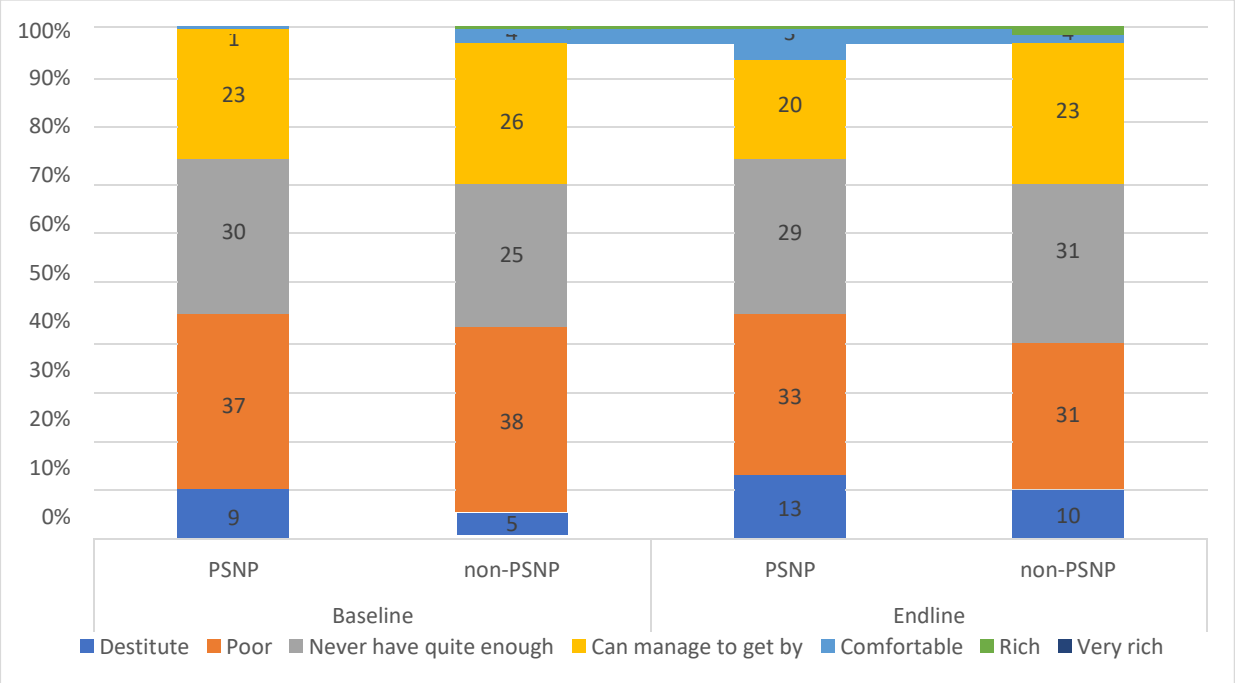


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

We also collected information on perception of absolute poverty which we refer to as subjective poverty/wellbeing. This is presented in Figure 7.2 by PSNP status and survey round. We observe that, at baseline, 46% of PSNP households consider themselves as destitute or poor. The corresponding

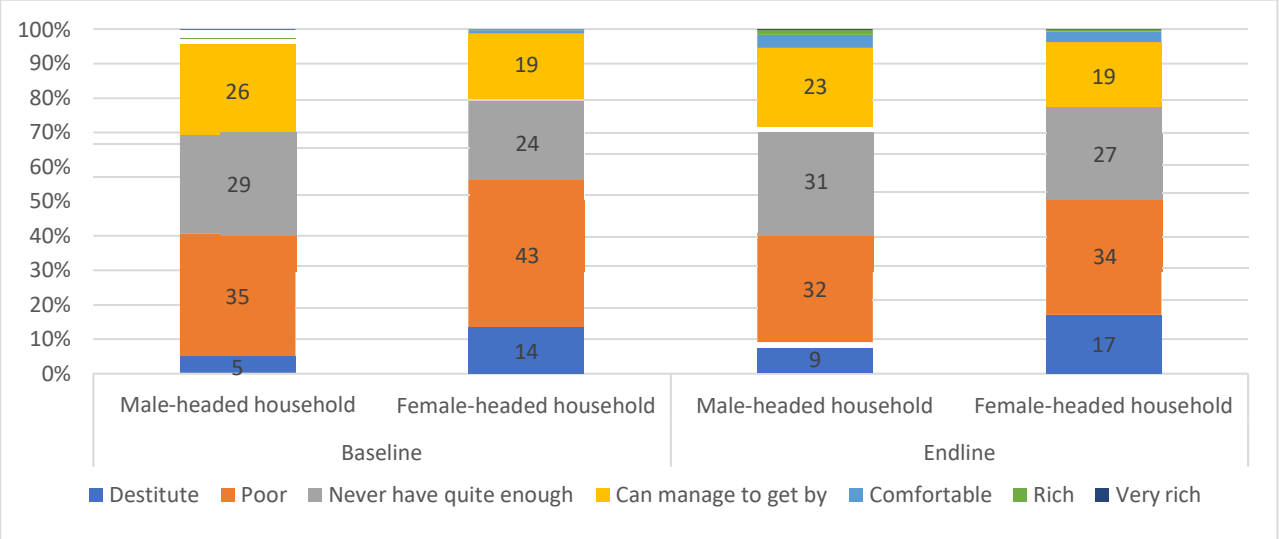
percentage among non-PSNP households is slightly lower at 43%. About a quarter of all households report that think they "can manage to get by". We do not see major changes over time in both groups. But do observe a deterioration among both groups. Among the PSNP households, the group of destitute households increased to 13% and it increased to 10% among the non-PSNP households.

Figure 7.2: Subjective poverty, by PSNP status and survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Figure 7.3: Subjective poverty, by headship and survey round

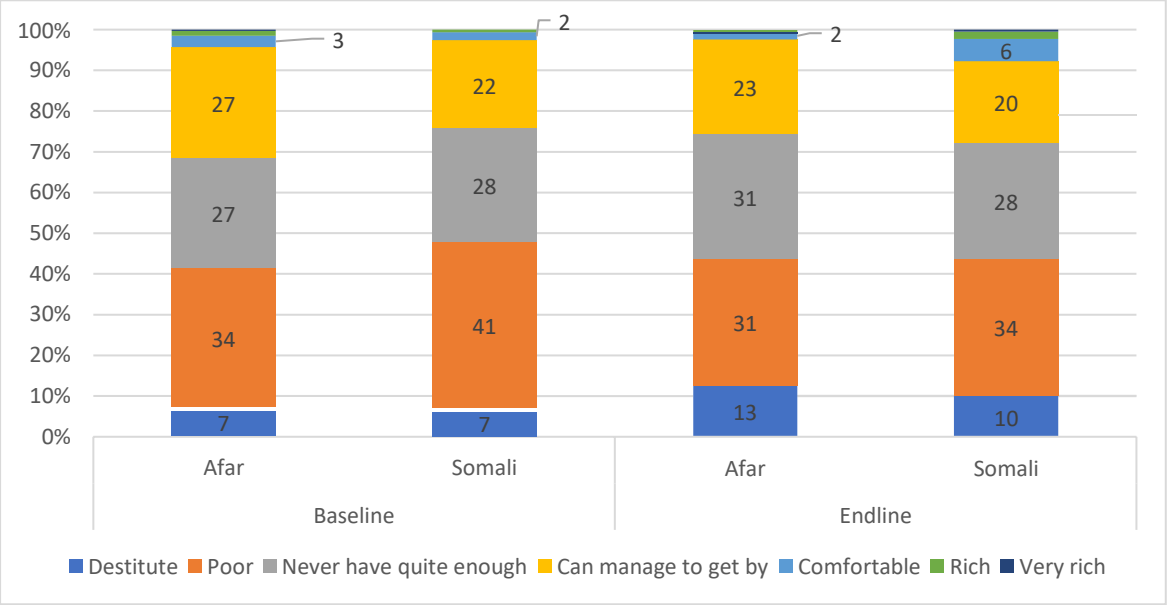


Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

Figures 7.3 and 7.4 present subjective poverty by sex of the household head and by region. Male-headed households are better off than female-headed household from the perspective of subjective poverty but over time both groups have seen some positive and negative changes. The two regions had similar levels of subjective poverty at baseline. Over time, the changes are mixed. For example, in

both regions the percentage of households that are destitute has increased. In Somali, the percentage of households that are comfortable has also increased.

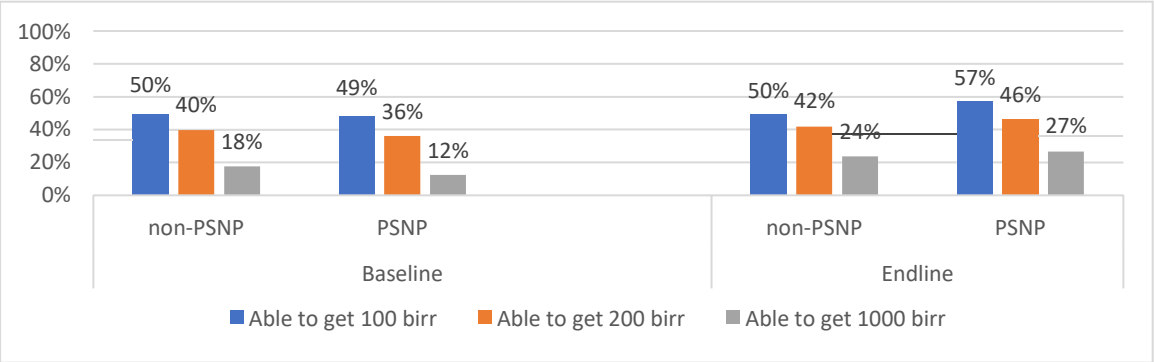
Figure 7.4: Subjective poverty, by region and survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Another important aspect of wellbeing and resilience is the ability to raise funds in case of an emergency. We asked, at baseline and at end-line, if the households needed a certain amount of money (100, 200 or 1000 birr) for an emergency could they obtain it within a week. Figures 7.5-7.7 present this by PSNP status, sex of the household head and by region. About half of the households (irrespective of PSNP status, at baseline, were able to obtain 100 birr. As the amount increases the proportion that can obtain it drops. Over time, we do not observe much change among the non-PSNP households. But among the PSNP households, the ability to obtain emergency funds has increased. At end-line, 57% of the PSNP households reported being able to obtain 100 birr and 27% were also able to raise 1000 birr.

Figure 7.5: Able to obtain emergency funds, by PSNP status and survey round

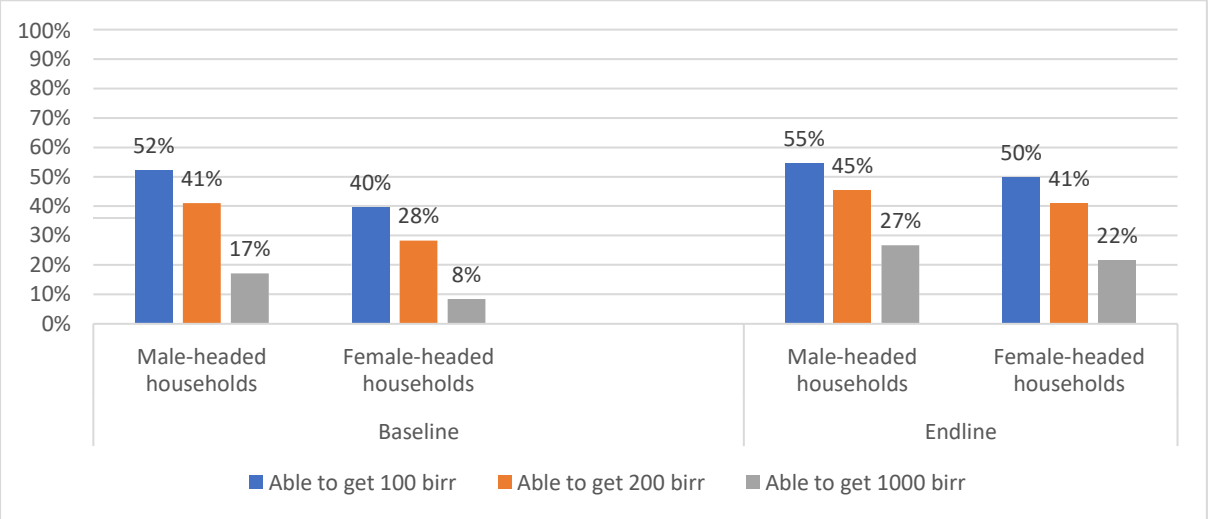


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Comparing male- and female-headed households, we find that a larger percentage of male-headed households are able to obtain emergency funds at baseline but, by end-line, both groups have seen

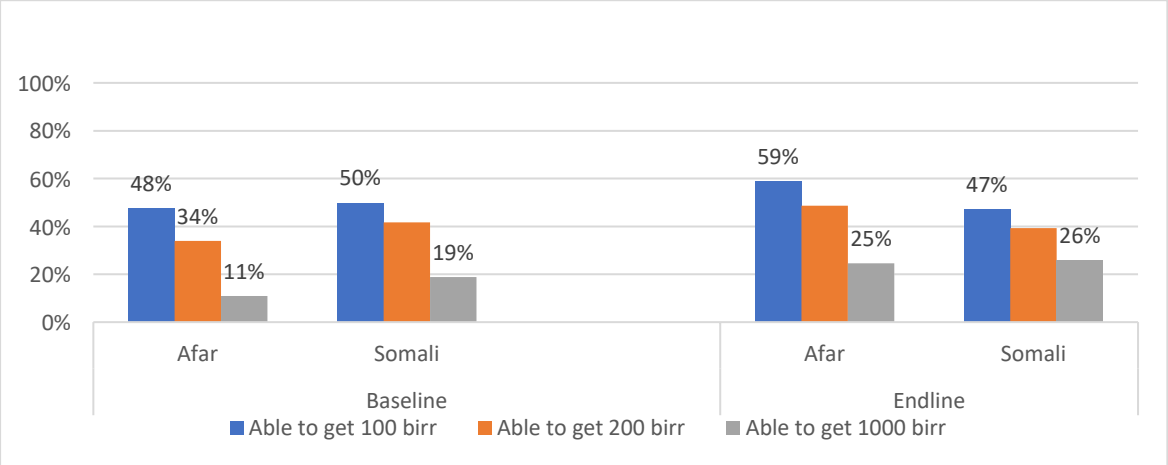
improvements in their ability to raise emergency funds and the gaps have significantly reduced (Figure 7.6).

Figure 7.6: Able to obtain emergency funds, by headship and survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

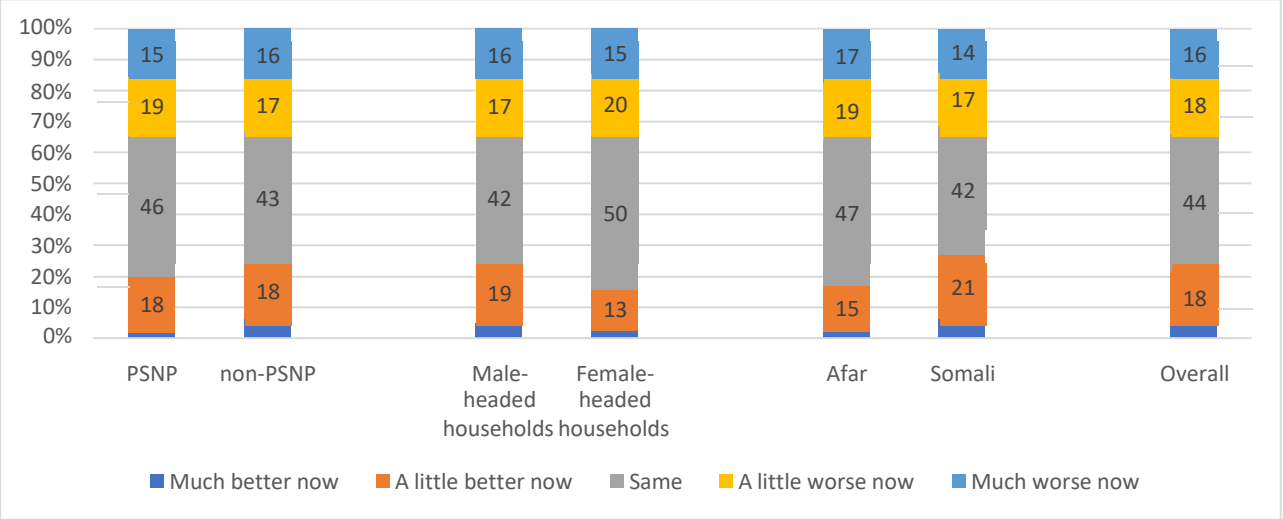
Figure 7.7: Able to obtain emergency funds, by region and survey round



Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

There is not much regional variation in the ability to raise emergency funds at baseline. By end-line, households in Afar show an improvement in their ability to raise emergency funds but we do not see such improvement in Somali (except in the case of accessing 1000Birr) (Figure 7.7).

Figure 7.8: Compared to two years ago, by PSNP status, headship and region at end-line



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

At end-line, we also asked households about their circumstances two years ago (Figure 7.8). 16% of the households overall reported doing much worse at end-line, 18% reported little worse, 44% reported doing same as two years ago, and 18% reported doing a little better. We do not observe significant variations across different groups.

While we observe some deterioration in relative economic standing and some positive and negative movements in subjective poverty. There is an improvement in the ability to obtain emergency funds over time. When households are asked to compare their circumstances to two years before, the large majority report either feeling the same or doing worse. In the next section we examine if the PSNP had an impact on subjective wellbeing.

7.3 Impact of the PSNP on subjective wellbeing

In this section, we present the impact estimate of participating in PSNP on subjective poverty. This indicator is derived from the same variable presented in figures 7.2-7.4 where each level is denoted by a number. For example, very rich gets a value of 7, rich gets a value of 6, comfortable gets a value of 5 and so on. Table 7.1 gives the estimates of the impact of PSNP on subjective poverty using the nearest neighbor matching method (described in chapter 2). We do not find any impact of the PSNP on subjective poverty.

Table 7.1: Impact on subjective poverty, Nearest Neighbor Matching

	Treatment effect (SE)	N	Interpretation
Change in subjective poverty	0.011 (0.093)	3381	We cannot reject the hypothesis that PSNP had no impact on this outcome. The estimated impact is not statistically different from zero.

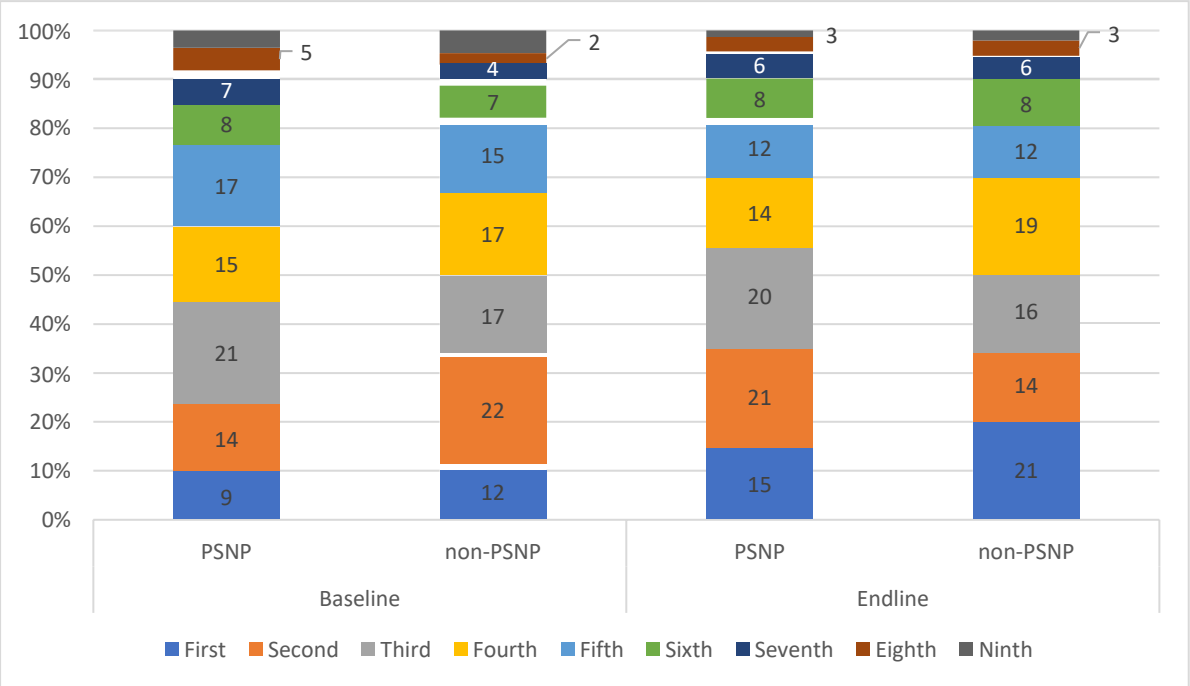
Note: The impact is estimated using a difference-in-difference approach combined with a kernel matching method. The standard errors (SE) are estimated using bootstrap methods with 500 repetitions. Asterisks indicate statistical significance: ***= p<0.01, **= p<0.05, and *= p<0.1.

Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

7.4 Agency

In this section, we present descriptive results on variables that reflect individuals' agency and comfort in speaking in public. We begin with a question, administered at baseline and end-line, which tells households to imagine a nine-step ladder, where on the bottom, the first step, are those who are totally unable to change their lives, while on step 9, the highest step, stand those who have full control over their own life. And then asks the step they are on. As seen in Figure 7.9, PSNP and non-PSNP households are very similar in this respect. If anything, we find that a greater percentage of non-PSNP households are on the bottom two steps than PSNP households (34% vs 23%). And, we observe some deterioration over time in both groups.

Figure 7.9: Control over life – 9-step ladder, by PSNP status and survey round

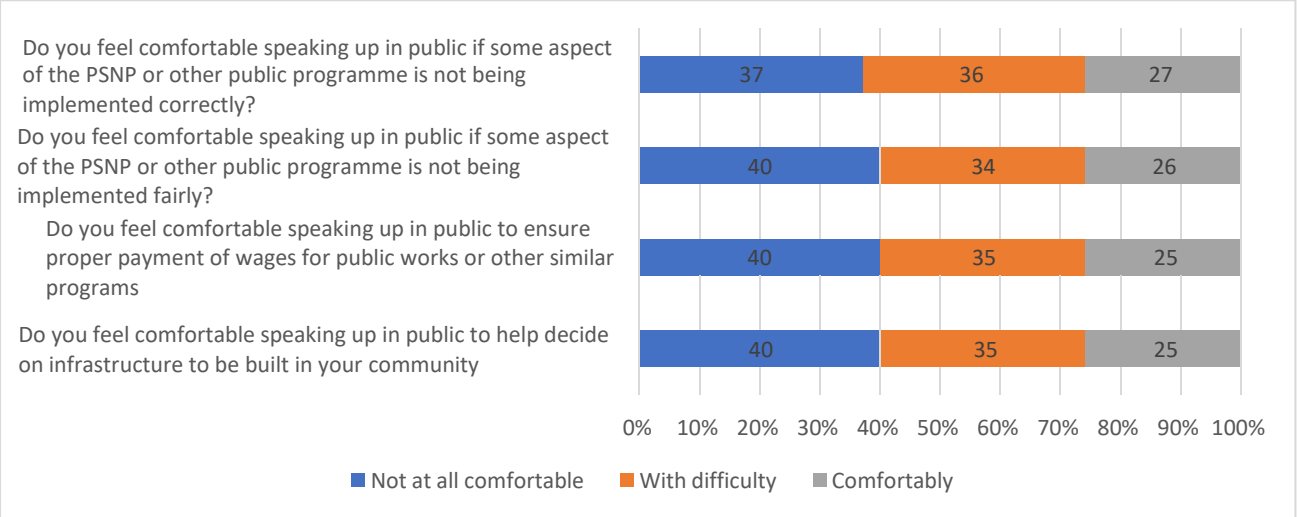


Source: Authors' computation using PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Figure 7.10 shows descriptive results for comfort in speaking up in public related to decisions regarding infrastructure, payments in the public works, if some aspects of the PSNP is not implemented fairly or correctly. We find that about 40% of the households are not at all comfortable speaking up in public.

Whereas 20-25% are comfortable. And the remaining 34-36% would speak up with difficulty. This reflects the head's comfort in speaking publicly.

Figure 7.10: Comfort in public speaking at end-line



Source: Authors' computation using PSNP4 End-line (2021) Household Survey.

7.5 Summary

In this chapter we presented results on subjective measures of wellbeing.

- PSNP household's perception of their economic standing is lower than that of non-PSNP households at baseline and at end-line. We observe some movement in this perception over time. We observe an increase in the percentage reporting "little poorer than most" and "the poorest" in both groups. Among the non-PSNP households, the percentage reporting themselves as "the poorest" has more than doubled during this time.
- In terms of absolute poverty, we observe that, at baseline, 46% of PSNP households consider themselves as destitute or poor. The corresponding percentage among non-PSNP households is slightly lower at 43%. Over time, among the PSNP households, the group of destitute households increased to 13% and it increased to 10% among the non-PSNP households.
- Male-headed households are better off than female-headed household from the perspective of subjective poverty but over time both groups have seen some positive and negative changes.
- The two regions had similar levels of subjective poverty at baseline. Over time, the changes are mixed.
- About half of the households (irrespective of PSNP status), at baseline, were able to obtain 100 birr. Over time, we observe small changes in this ability among the non-PSNP households. But among the PSNP households, the ability to obtain emergency funds has increased appreciably.
- Comparing male- and female-headed households, we find that a larger percentage of male-headed households are able to obtain emergency funds at baseline but, by end-line, both

groups have seen improvements in their ability to raise emergency funds and the gaps have significantly reduced.

- There is not much regional variation in the ability to raise emergency funds at baseline. By end-line, households in Afar show an improvement in their ability to raise emergency funds but we do not see such improvement in Somali.
- Compared to two years ago, 16% of the households overall reported doing much worse at end-line, 18% reported little worse, 44% reported doing same, and 18% reported doing a little better. We do not observe significant variations across different groups.

The research question we set out to address in this chapter was **RQ3: Has PSNP reduced poverty?**

- We can only partially answer this question as we did not collect consumption expenditure data in the lowlands based on CSA's past experience of difficulty in obtaining reliable consumption data using standard household surveys.
- We do not find any impact of the PSNP on subjective poverty

This chapter also presents descriptive statistics on agency.

- PSNP and non-PSNP households are very similar in their perception of the extent to which they have full control over their life. Over time, we observe some deterioration in this perceived extent in both groups.
- In terms of speaking up in public – on matters relating to infrastructure, payments in the public works, if some aspects of the PSNP is not implemented fairly or correctly – more than 40% of the respondents are not at all comfortable, about 20-25% are comfortable and the remaining 34-36% would speak up with difficulty.

Appendix A: Sampling Strategy

Sampling for this phase-4 evaluation round is more complicated for a number of reasons. We outlined these complications in the inception report and the final sampling strategy was discussed with and endorsed by the DCT as well as participants of the inception workshop in December 2015.

A.1 Determining sample size

The size of the sample depends on a number of considerations, a lot of which have been noted in the previous sub-section. The next paragraphs summarize and highlight these as required.

Purpose of the survey: The survey is expected to generate baseline information necessary to monitor performance and outcome indicators of PSNP4 as well as evaluate the programme's impact at the regional as well as the national level. Recall that the survey will be implemented in six regions: Tigray, Amhara, Oromia, SNNP ("Highlands") and Afar and Somali ("Lowlands"). Both beneficiaries and non-beneficiaries are selected from PSNP4 woredas and kebeles.

Primary indicator: The size of the sample is in part determined by indicators being considered primary for PSNP4. Indicators characterized by high levels of variability demand larger sample sizes to fully capture their distribution. The food gap is identified as the primary indicator for the evaluation.

Significance and power: The sample has to be sufficiently large to minimize the chance of detecting an effect that does not exist (statistical significance) and maximize the chance of detecting an effect that does exist (statistical power). Following standard practice, these will be set at a target level of significance of 5% (two-tailed) and statistical power of 80%.

Minimum detectable effect size: Sample size depends on the minimum level of impact (known as minimum detectable effect sizes) the survey is desired to detect in the relevant indicator. For example, should the sample size be large enough to detect that PSNP4 transfers have reduced the food gap by 0.25 months, or 0.5 months, or by 0.75 months? Smaller effect sizes require larger samples; conversely, larger effect sizes require smaller samples.

Design effect: The design effect reflects the extent to which the indicator of choice is correlated across households or individuals within a specified group or cluster, usually defined by geographic location.[7] Intra-cluster correlations (ICCs) and cluster-level smaple size are used to measure this extent. Higher ICCs mean that the design effect is stronger and that larger samples are needed.

Attrition: We need to take into account the fact that over time some households will move to other localities, others break-up with members dispersing, still others may chose not to continue to be interviewed. Based on our experiences with other longitudinal household surveys in rural Ethiopia, we assume that ten per cent of the sample will attrit between baseline and end-line (in five years and over three rounds).

Table A1: Estimated and selected parameters for sample size determination

[7] More formally, the design effect is the ratio between the variance (and thus the required sample size) associated with complex sample design (cluster or multistage sampling) and the variance (or sample size) if the sample had been drawn using simple random sampling of the ultimate respondents.

Variable	Level	
	Highlands	Afar and Somali
Statistical significance (two-tailed)	5%	5%
Statistical power	80%	80%
Proportion of program beneficiaries in the sample [a]	50%	50%
Intra-cluster correlation	0.14	0.14
Cluster sample size (per Enumeration Area (EA))	28	30
Design effect	5.3	5.6
Attrition (across three rounds over five years)	10%	10%
Minimum detectable effect size (MDE)[b]	**0.3SD=0.58 months reduction in food gap**	0.4SD=0.49 months reduction in food gap

Notes: [a] The share of beneficiaries has been tweaked slightly and is a bit higher than half in the actual sample. [b] MDEs are computed as a fraction of the standard deviation of the distribution (SD) of the food gap. Both this fraction and the corresponding absolute level of the food gap are reported.

We started with data on the distribution of the food gap in the PSNP woredas. These data were generated by the five rounds of surveys linked to the evaluation of PSNP and collected by the CSA over the last 10 years. The average ICCs at the kebele level (strictly speaking, at the Enumeration Area (EA) level) were computed for the Highlands and the Lowlands. These were subsequently used to compute the applicable design effect. It turned out that the ICC averged around 0.14 for both groups and the corresponding design effect comes out as 5.3 (Highlands) and 5.6 (Lowlands). The number of household to be sampled per EA, respectively 28 and 30, explain the later difference.

Table A1 above summarizes the estimated and selected parameters for sample size determination. Based on these parameters, the size and composition of the sample are calculated. The results are reported in Tables A2-3.

Table A2: Sample Sizes

Region	**Sample size**	**Number of woredas**	**Number of EAs / kebeles**
Tigray, Amhara, Oromia, SNNP (*Highlands*)	1850 (7400)	22 (88)	66 (264)
Afar, Somali (*Lowlands*)	1080 (2160)	12 (24)	36 (72)

Note: Figures in brackets are the corresponding totals for the Highlands and Lowlands, respectively.

Table A3: Sample Composition in each EA or kebele

Region	**Sample Composition in each EA or kebele:**		
	Public Works Beneficiaries	**Permenant Direct Support Beneficiaries**	**Non-beneficiaries**
Tigray, Amhara, Oromia, SNNP	13	3	12
Afar, Somali	13	4	13

The total size of the sample needed is 9,560 households across the six regions. Each region in the Highlands should have a sample of 1,850 household over 22 woredas and 66 EAs. This

adds up to 7,400 households in 264 EAs across 88 woredas. The two regions in the Lowlands group equally share 2,160 households across 24 woredas and 72 EAs (or kebeles).

A.2 Panel surveys vs. repeated cross-sections

Here we note two points. First, we have worked extensively with CSA on implementing household panel surveys with particular attention being paid to survey protocols that minimize sample attrition. For example, the attrition rate between 2006 and 2014 was 1.9 percent per year, a rate lower than that found in panel surveys such as the highly regarded US Panel Survey of Income Dynamics.

Second, the evaluation of PSNP3 used a panel design for the Highlands and a repeated cross-section design for the Lowlands. The repeated cross-section design was used over concern that it would be difficult to track pastoralist households over time. We followed this strategy in this PSNP4 evaluation; a panel of households is followed in the Highlands over three survey rounds (2016, 2018, 2021) and three repeated cross-sectional surveys (2016, 2018, 2021) are conducted in the Lowlands.

A.3 Sample selection

Three steps were involved in the selection of households for the PSNP-4 baseline. First, the 112 woredas were randomly selected from the pool of PSNP-4 woredas using proportions derived from population size and project coverage. At the second stage, 3 EAs were randomly chosen among EAs in each woreda. The final step was the selection of households from within each EA (28 and 30, respectively, in the Highlands and Lowlands). This was done based on a fresh listing of households residing within each EA. The listing form used for this purpose gathers information on household current and past PSNP beneficiary status; age and gender of the household head; household land and livestock holdings; and household wealth self-ranking relative to other village residents[8]. Households were randomly selected from this list until the desired number and composition of households was obtained (see Table A4). To maximize the chance of obtaining a control sample that is as similar as possible to the treatment sample, the non-beneficiary (control) households were chosen from the bottom four rungs of the subjective wealth ranking.

Table A4: Sample Composition in each EA or kebele

	Sample Composition in each EA or kebele		
	Public Works Beneficiaries	**Permenant Direct Support Beneficiaries**	**Non-beneficiaries**
Tigray, Amhara, Oromia, SNNP	13	3	12
Afar, Somali	13	4	13

Finally, Table A5 shows how the sample in the highlands (panel) and lowlands (repeated cross-section) evolved over time.

Table A5: Sample dynamics

[8] During listing, we asked all households to place themselves on to a poverty ladder that has 7 rungs. The first rung represented the poorest households in the village and the highest (7th) rung the richest households in the village.

Highlands (panel of households) excluding the Tigray sub-sample	
Target sample:	5,544 households in 66 woredas and 198 EAs
Achieved sample at the baseline:	5,493 in 66 woredas and 198 EAs
Achieved sample at the midline:	5,271 in 66 woredas and 197 EAs
Achieved sample at the end-line:	5,111 in 66 woredas and 196 EAs
Attrition/non-response rates:	
Target sample vs achieved baseline sample:	0.93 %
Achieved baseline sample vs achieved midline sample:	4.21 %
Achieved baseline sample vs end-line sample:	8.47 %
Lowlands (repeated cross-section)	
Target sample:	2,160 households in 24 woredas and 72 EAs
Achieved sample at the baseline:	1,983 households in 23 woredas and 70 EAs
Achieved sample at the midline:	1,945 households in 21 woredas and 65 EAs
Achieved sample at the end-line:	2,084 households in 22 woredas and 70 EAs
Attrition/non-response rates:	
Target sample vs achieved baseline sample:	8.93 %
Target sample vs achieved midline sample:	10.77 %
Target sample vs achieved end-line sample:	3.65%

Source: Authors' computations using data form the PSNP4 Baseline (2016), Midline (2018), and End-line (2021) Household Surveys.

Appendix B: Response to Comments on Factors Potentially Conditioning Impact

In this Appendix, we document responses to comments received regarding other factors potentially conditioning impact and how some of them could be brought into the analysis, through heterogeneity analysis, for example.

1. Factors potentially conditioning impact

   Comment: As per the ToRs "The analysis of RQs 1 to 4 will be extended to determine whether observed impacts were affected by: Location (Highlands, Lowlands); Initial levels of food insecurity; Initial levels of wealth; gender of household head; participation in Public Works and Direct Support." Please add the impact disaggregated as per the ToRs. Also note that the ToRs state that "All questions included in this ToR should be addressed by the organization. Any statistical power issues, potential bias, or other caveats of the assessment will need to be reported accordingly when presenting the results."

   Elsewhere, the comments request to examine how impacts differ by the duration of program participation, by predictability of transfers, modality of transfers and by transfer adequacy.

   Response: We provide a response for each of the requested pieces of analysis.

   i. Location (Highlands, Lowlands): We address this request fully by having separate outcomes reports for the Highlands and Lowlands.

   ii. Initial levels of food insecurity: We are adding new impact estimates to the Highlands Outcomes Report to differentiate impacts by initial level of food insecurity, using the baseline food gap.

   iii. Initial levels of wealth: We are adding new impact estimates to the Highlands Outcomes Report to differentiate impacts by initial level of wealth, measured by baseline livestock holdings (in TLUs).

   iv. Gender of household head: We find that only 29% of PSNP households and 21% of non-PSNP households are female headed. It will not be feasible to conduct a robust matching model in a sample this small based on our experience with the matching models in the PSNP4 data. However, the estimates in the report do control for gender of the household head.

   v. Participation in Public Works and Direct Support: Following the principles of 'one-PSNP', we have devised a method to estimate the average impact of the whole program: Public Works (PW) and Direct Support (DS). As described in the Performance Report for the end-line survey, the PW and DS programs operated similarly in terms of timing and rate of payments. Moreover, in some households pregnant women transitioned from PW to Temporary DS under PSNP4. As a result, it is possible to treat PW and DS as comparable parts of a unified PSNP and pool the programs for the impact estimates, and it would not be possible to accurately isolate the effects of PW from DS. This similarity in PW and DS is also consistent with a key assumption needed for these matching models to provide unbiased causal estimates of impact, the Stable Unit Treatment Value Assumption

(SUTVA). SUTVA requires that the treatment status of one unit does not affect potential outcomes of any other unit (a non-interference assumption) and that treatments are constant for all units. Finally, only around 20% of households receive DS, which is again too small of a sample to allow estimation of robust matching models in these data.

vi. Duration of program participation: We have investigated the patterns of program participation for the performance report and found that there is not enough variation in program participation to be able to estimate matching models that would be capable of isolating the effect of differences in duration of program participation. Sections 4.2-4.3 of the Performance Report discuss the relevant issue. For the highlands, it states that:

- ". (a) conditional on selection into the PSNP, participation was constant for Highlands households between 2016 and 2021 (71 percent were PSNP participants for five or six years over this time period); (b) households that were included for five years out of six were usually excluded in 2016 but included thereafter; (c) households that were included for only one or two years were usually included in 2016 and 2017 but excluded after that. This suggests a pattern whereby there was some movement in and out of the program in the first years of PSNP4 but after 2017, participation (or non-participation) was constant. This small amount of re-targeting is consistent with what regional, woreda and kebele officials told us (section 4.2."

- The same assessment covering 2016-2021 cannot be made for the Lowlands since the Lowlands' sample is not a panel. Nevertheless, an analogous analysis covering 2018-2021 can be conducted using the end-line survey data alone. The results are comparable to what was found for the Highlands.

  Thus, there is an insufficient sample with low participation for a credible analysis of the role of duration.

vii. Predictability of transfers: a major purpose of the outcomes reports is to measure the average impact of PSNP4 transfers on household wellbeing, which captures the average effect of delayed transfers. However, disaggregating the analysis by the timing of when transfers are received is likely to provide estimates with low power. Those estimates may also be biased since predictability of transfers is likely to be determined primarily by unobserved local factors.

viii. Modality of transfers: Transfer modalities (food or cash) are largely determined by region, with food being the primary modality of transfer in Afar, where the data is relatively sparse. Other regions provide beneficiaries with a blend of food and cash. Section 6.6 of the Performance Report notes that:

> "Our payments data show that the use of these modalities differs sharply by region. For example, between Tir and Miazia EC2012, the percentages of PSNP clients paid *only* in cash were 99, 94 and 78 in SNNP, Somali and Amhara respectively. In the same time period, 82 percent of PSNP clients in Afar were *only* paid in in-kind. We cannot meaningfully compare the frequency of cash and in-kind payments because we cannot tell whether any

differences reflect something specific to the transfer modality or something specific to the region where it is used."

Therefore, it is not possible to estimate how impacts differ by payment modality, or to separate the effect of modality from regional effects.

ix. Transfer adequacy: the size of transfer received is likely to be correlated with both observable and unobservable characteristics, so measuring the effect of different amounts of transfers received is likely to be confounded with other measures of household characteristics. It would not be possible to adequately control for this in the impact estimates.

Appendix C: Additional Tables - Chapter 7

Table 7A.1: Self-perceived economic standing in village, by PSNP status and survey round

Baseline

	PSNP		Non-PSNP		p-value of difference by PNSP status	Male-headed		Female-headed		p-value of difference by gender of HHH
	Mean	N	Mean	N		Mean	N	Mean	N	
The richest in the village	0.000	1,035	0.007	938	0.005	0.004	1,459	0.002	514	0.478
Amongst the richest in the village	0.008	1,035	0.019	938	0.026	0.015	1,459	0.008	514	0.212
Richer than most households	0.038	1,035	0.049	938	0.215	0.051	1,459	0.021	514	0.005
About average	0.328	1,035	0.355	938	0.199	0.373	1,459	0.249	514	0.000
A little poorer than most households	0.231	1,035	0.210	938	0.264	0.224	1,459	0.212	514	0.571
Amongst the poorest in the village	0.300	1,035	0.294	938	0.762	0.276	1,459	0.358	514	0.001
The poorest in the village	0.096	1,035	0.065	938	0.013	0.057	1,459	0.150	514	0.000

End-line

	PSNP		Non-PSNP		p-value of difference by PNSP status	Male-headed		Female-headed		p-value of difference by gender of HHH
	Mean	N	Mean	N		Mean	N	Mean	N	
The richest in the village	0.007	1,046	0.006	1,038	0.792	0.007	1,494	0.005	590	0.674
Amongst the richest in the village	0.011	1,046	0.013	1,038	0.679	0.014	1,494	0.008	590	0.301
Richer than most households	0.035	1,046	0.054	1,038	0.040	0.046	1,494	0.042	590	0.754
About average	0.279	1,046	0.294	1,038	0.459	0.311	1,494	0.224	590	0.000
A little poorer than most households	0.261	1,046	0.243	1,038	0.338	0.257	1,494	0.239	590	0.393
Amongst the poorest in the village	0.287	1,046	0.239	1,038	0.013	0.241	1,494	0.319	590	0.000
The poorest in the village	0.120	1,046	0.151	1,038	0.034	0.124	1,494	0.163	590	0.022

Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Household Surveys.
Notes: p-values are from a t-test for of equality of means of the outcome across the sub-samples.

Table 7A.2: Subjective poverty, by PSNP status and survey round

Baseline

	PSNP		Non-PSNP		p-value of difference by PNSP status	Male-headed		Female-headed		p-value of difference by gender of HHH
	Mean	N	Mean	N		Mean	N	Mean	N	
Destitute	0.091	1,032	0.053	936	0.001	0.051	1,457	0.135	511	0.000
Poor	0.368	1,032	0.379	936	0.613	0.355	1,457	0.427	511	0.004
Never have quite enough	0.296	1,032	0.251	936	0.027	0.288	1,457	0.237	511	0.027
Can manage to get by	0.229	1,032	0.261	936	0.099	0.263	1,457	0.190	511	0.001
Comfortable	0.013	1,032	0.037	936	0.000	0.030	1,457	0.008	511	0.005
Rich	0.004	1,032	0.015	936	0.010	0.011	1,457	0.004	511	0.149
Very rich	0.000	1,032	0.003	936	0.069	0.002	1,457	0.000	511	0.305

End-line

	PSNP		Non-PSNP		p-value of difference by PNSP status	Male-headed		Female-headed		p-value of difference by gender of HHH
	Mean	N	Mean	N		Mean	N	Mean	N	
Destitute	0.129	1,038	0.100	1,021	0.038	0.092	1,473	0.171	586	0.000
Poor	0.331	1,038	0.313	1,021	0.383	0.317	1,473	0.336	586	0.402
Never have quite enough	0.287	1,038	0.307	1,021	0.334	0.308	1,473	0.270	586	0.089
Can manage to get by	0.202	1,038	0.231	1,021	0.112	0.229	1,473	0.186	586	0.034
Comfortable	0.033	1,038	0.038	1,021	0.505	0.037	1,473	0.031	586	0.464
Rich	0.013	1,038	0.010	1,021	0.556	0.014	1,473	0.005	586	0.100
Very rich	0.005	1,038	0.001	1,021	0.106	0.003	1,473	0.002	586	0.522

Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Household Surveys.

Notes: p-values are from a t-test for of equality of means of the outcome across the sub-samples.

Table 7A.3: Summary statistics of emergency

	Baseline									
	PSNP		Non-PSNP		Difference by PNSP status (p-value)	Male-headed		Female-headed		Difference by gender of HHH (p-value)
	Mean	N	Mean	N		Mean	N	Mean	N	
Able to get 100 Birr for an emergency	0.485	1,035	0.496	938	0.635	0.521	1,465	0.395	516	0.000
Able to get 200 Birr for an emergency	0.362	1,035	0.397	938	0.117	0.410	1,465	0.283	516	0.000
Able to get 1000 Birr for an emergency	0.124	1,035	0.176	938	0.001	0.171	1,465	0.083	516	0.000
	End-line									
	PSNP		Non-PSNP		Difference by PNSP status (p-value	Male-headed		Female-headed		Difference by gender of HHH (p-value)
	Mean	N	Mean	N		Mean		Mean	N	
Able to get 100 Birr for an emergency	0.570	1,046	0.496	1,038	0.001	0.546	1,494	0.500	590	0.057
Able to get 200 Birr for an emergency	0.465	1,046	0.419	1,038	0.036	0.454	1,494	0.410	590	0.067
Able to get 1000 Birr for an emergency	0.267	1,046	0.237	1,038	0.118	0.266	1,494	0.215	590	0.015

Source: Authors' computations using data form the PSNP4 Baseline (2016) and End-line (2021) Household Surveys.
Notes: p-values are from a t-test for of equality of means of the outcome across the sub-samples.

Appendix D: Mid-upper Arm Circumference Among Children and Women

The results regarding mid-upper arm circumference (MUAC) among children under 5 and women under 40 in the lowlands sample are reported in Table D.1. The mean MUAC among children is 13 and the mean MUAC among women is 23.5. About 13% and 17% children under 5 years of age among PSNP households and non-PSNP households, respectively, are categorized as moderately acute malnourished (MAM) (MUAC between 11-12.5cm). The difference between the two groups is statistically significant. Furthermore, about 14% and 15% children under 5 are identified as severe acute malnourished (SAM) (MUAC less than 11cm) among PSNP and non-PSNP households, respectively, though this difference is not statistically significant. Similarly, there is no statistically significant difference between PSNP and non-PSNP households in terms of Mothers' average MUAC.

Table D.1: MUAC among children and mothers by the Treatment and Control samples in 2021

	PSNP households			Non-PSNP households			p-value
	Mean	SD	N	Mean	SD	N	
Child average MUAC	13.082	2.837	774	13.093	3.368	912	0.939
% Children suffer from Moderate Acute Malnutrition	0.131	0.337	995	0.166	0.372	1150	0.021
% Children suffer from Severe Acute Malnutrition	0.137	0.344	995	0.151	0.359	1150	0.337
Mother average MUAC	23.494	3.627	875	23.375	3.774	1007	0.485

Notes: SD refers to standard deviations. p-values are from a t-test for of equality of means of the outcome across the PSNP and non-PSNP samples.

Source: Authors' computations using data form the PSNP4 End-line (2021) Community Survey.

Appendix E: Health Services at the Community Level

In this appendix, we report descriptive results on health services at the community level. 38% of the kebeles reported that there is a government health center in their kebele and 86% reported that there is a government health post. 33% and 89% of the kebeles in Afar and 43% and 83% in Somali reported there being a government health center and government health post, respectively. There are 2.2 health extension workers at each health post. A fraction of the kebeles report undertaking nutrition activities. 69%, 70% and 66% report undertaking community-based nutrition activities, growth monitoring and therapeutic feeding, respectively. They also report conducting IYCF counseling at the health post or at a home visit and food demonstrations at least once in the last 3 months.

There were questions about their understanding of PSNP's co-responsibilities associated with nutrition. On average, 3.2 antenatal visits, 2.9 nutrition BCC sessions and 3.5 postpartum visits for pregnant women receiving Temporary Direct Support were reported. 70% and 82% reported that lactating women (who are receiving Temporary Direct Support) with a child less than one year old attend growth monitoring and promotion/behavioural change communication sessions and ensure that child receives routine immunizations, respectively.

When asked about what could constitute as contributions to Public Works' clients PW commitment – 66% reported participation in up to two community health days, 70% reported attending nutrition BCC sessions provided at PW site, 70% reported participation in community conversations as part of the community-based nutrition activities and 73% reported looking after children at childcare center located at a PSNP worksite.

There are about 4 community health workers or members of the Health Development Army. 23% and 44% kebeles report that there is a social worker and a community care coalition operates, respectively.

Table E.1: Health services at community level, by region

	All			Afar			Somali		
	Obs.	Mean	SD	Obs.	Mean	SD	Obs.	Mean	SD
% Kebeles have a government health center	71	38%	49%	36	33%	48%	35	43%	50%
Distance to closest health center outside the kebele	44	12.09	12.60	24	11.00	14.73	20	13.40	9.66
% Kebeles have a government health post	71	86%	35%	36	89%	32%	35	83%	38%
Distance to closest health post outside the kebele	10	2.90	2.23	4	3.25	2.50	6	2.67	2.25
No. of health extension workers at this post	71	2.20	1.26	36	2.25	1.25	35	2.14	1.29
% Health posts undertaking community-based nutrition activities	71	69%	47%	36	72%	45%	35	66%	48%
% Health posts provide growth monitoring	71	70%	46%	36	72%	45%	35	69%	47%
% Health posts provide therapeutic feeding	71	66%	48%	36	81%	40%	35	51%	51%
No. of times IYCF counselling at health post were conducted, last 3 months	71	1.76	2.17	36	2.33	2.47	35	1.17	1.64
No. of times IYCF counselling at home visit were conducted, last 3 months	71	1.52	1.94	36	1.97	2.26	35	1.06	1.43
No. of times community conversations about IYCF were conducted, last 3 months	71	1.39	1.95	36	2.00	2.41	35	0.77	1.03
No. of times food demonstrations were conducted, last 3 months	71	1.13	1.89	36	1.69	2.20	35	0.54	1.29
No. of ante-natal consultations should pregnant women getting Temporary Direct Support receive	53	3.32	1.85	28	3.96	1.43	25	2.60	2.02
No. of nutrition behavior change communication (BCC) sessions should pregnant women receiving Temporary direct support receive	41	2.98	1.62	26	3.31	1.57	15	2.40	1.59
No. of post-partum health facility visits lactating women with a child less than one year old should attend if they are receiving Temporary Direct Support	44	3.50	1.70	29	3.72	1.75	15	3.07	1.58

% reported that lactating women with a child less than one year old attend growth monitoring and promotion/behavioral change communication sessions if they are receiving Temporary Direct Support	71	70%	46%	36	86%	35%	35	54%	51%
% reported that lactating women with a child less than one year old who are receiving Temporary Direct Support must ensure that the child receives routine immunizations	71	82%	39%	36	97%	17%	35	66%	48%
Activity can be counted as contribution towards Public Works' clients Public Works commitment									
Participation in up to two community health days	71	66%	48%	36	81%	40%	35	51%	51%
Nutrition Behavior Change Communication (BCC) sessions provided at a Public Works site	71	70%	46%	36	89%	32%	35	51%	51%
Participation in community conversations conducted as part of community-based nutrition	71	70%	46%	36	89%	32%	35	51%	51%
Looking after children at a childcare center located at a PSNP worksite.	71	73%	45%	36	86%	35%	35	60%	50%
No. Village Community Health Workers or members of the Health Development Army in the kebele	71	3.96	6.02	36	4.56	5.83	35	3.34	6.24
% Kebeles have social worker	71	23%	42%	36	19%	40%	35	26%	44%
% Kebeles where community care coalition operates	71	44%	50%	36	44%	50%	35	43%	50%
Distance from kebele to the nearest primary hospital (km)	71	46.70	51.57	36	49.56	58.57	35	43.77	43.90
Distance from kebele to the nearest referral hospital (km)	71	137.44	130.92	36	158.22	126.70	35	116.06	133.54

Notes: 'Obs.' and 'SD' stand for 'number of observations' and 'standard deviation', respectively.
Source: Authors' computations using data form the PSNP4 End-line (2021) Community Survey.