# The Productive Safety Net Programme IV

# End-line Outcomes Report (2021) – Highlands

[a] International Food Policy Research Institute

[b] Cornell University

February 4, 2022

# Table of Contents

# List of Tables

## List of Figures

# Executive Summary

This report documents the impact of the Fourth Phase of the Productive Safety Nets Programme (PSNP) from 2016–2021 in the Highlands regions of Ethiopia: Amhara, Oromia, and SNNP. This End-line Outcomes Report for the Highlands builds on evidence presented in the 2018 Midline Outcomes Report. It presents the end-line assessments of program impact and addresses the outcome-related research questions listed in the terms of reference for this study. This Executive Summary is organized as a set of answers to these research questions, based on the evidence presented in the report.

***RQ17: How are transfers used?***

- The end-line survey data show that over 90 percent of the households in highland areas spend cash transfers to buy food and about 44 percent to buy nonfood items.
- The results also indicate that households use nearly three-quarters of cash transfers to buy food and about 17 percent to buy nonfood items.
- About 67 percent of the food transfer is in storage and 12 percent has been sold. Nearly 18 percent was given to other households which, compared to the 3 percent cash transfers given to other households, indicates that households are more likely to provide in-kind assistance than cash assistance to other households.

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

- The average PSNP household reports higher levels of food gap than non-PSNP households both at baseline and end-line. However, between baseline and end-line, the mean food gap decreased by 0.2 months for PSNP households but increased by 0.1 months for non-PSNP households.
- Consistent with previous assessments on the same sample, the household dietary diversity score (HDDS) is dismally low among households in the sample and, on average, households in these areas consume from only 4.5 food groups out of the possible 12 food groups at end-line. This improved from baseline by an average of an increase of 0.4 food groups for all households.
- Overall, between baseline and end-line, mean per capita monthly total expenditures for all households increased by 33 percent. This increase comes from the increase in per capita monthly food expenditures, which rose by 39 percent over the same period. Per capita monthly nonfood expenditures, on the other hand, declined by 4 percent.
- Negligible difference exists between PSNP and non-PSNP households in terms of the mean growth rate in real per capita consumption expenditures between 2016 and 2021. In terms of levels, PSNP households have slightly lower real per capita monthly expenditures than non-PSNP households in both rounds. We note that these latter mean differences between PSNP and non-PSNP households are statistically significant at the 95 percent level.

- Results from the impact analysis show that the PSNP was able to reduce the food gap, on average, by about six days per year but did not have a statistically significant impact on household dietary diversity or on changes in real per capita monthly expenditures.

***RQ2: Has PSNP impacted household resilience to shocks?***

- On average, between baseline and end-line, livestock ownership of PSNP households increased by about 0.5 Tropical Livestock Units (TLU) (compared to non-PSNP households, where ownership grew by 0.2 TLU on average).
- The impact analysis show that the PSNP increased livestock TLU by 25 percent – this impact is statistically significant at the 5 percent significance level.
- Moreover, the average value of livestock assets more than doubled in the two periods (although this change does not appear to be associated with the PSNP and shows a counterintuitive result, perhaps due to measurement error).
- Droughts remain the most important shocks among PSNP and non-PSNP households, followed by floods and erosion. Drought shock is reported highest among households in Amhara, followed by Oromia and SNNP. Overall, about 15 percent of PSNP (13 percent of non-PSNP) households from Amhara report drought as the first most important shock. PSNP households are more likely to report experience of shocks than non-PSNP households. While the PSNP is designed to mitigate some of these shocks, the evidence from chapter 6 of the Performance Report shows that payments were neither predictable nor made in a timely fashion or in full entitlements.
- In line with this, reports of distress asset sales are widespread. Increasing trends are observed over time in the proportion of households reporting distress sales of assets. The most reported distress asset sales are of livestock assets for food and emergency cash needs. However, PSNP households are more likely to report these sales across all regions. Results from the impact analysis show that the PSNP4 did not have an impact on reducing distress sales of assets.

***RQ3: Has PSNP reduced poverty?***

- We measured poverty using objective as well as subjective measures of wellbeing.
- Starting with relative subjective wellbeing, PSNP households' perception of their economic standing is lower than that of non-PSNP households at both baseline and end-line. We do not observe much movement in this perception over time.
- In terms of perception of absolute poverty, at baseline, a large percentage of PSNP households (64 percent) consider themselves as destitute or poor. The corresponding percentage among non-PSNP households is lower, at 41 percent. About one-quarter of non-PSNP households that think they "can manage to get by," whereas only 9 percent of PSNP households report this. We see some improvements over time in both groups.

- At baseline, 37 percent of PNSP households and 52 percent of non-PSNP households report they would be able to obtain 100 Birr for an emergency. As the amount increases, the proportion that can obtain it drops. Over time, we observe a considerable improvement in PSNP households' ability to obtain such emergency funds.
- Almost one-quarter (23 percent) of households overall report doing much worse now compared to two years ago; 20 percent report doing a little worse now; 35 percent report doing the same as two years ago; and 19 percent report doing a little better now. We do not observe significant variations across different groups.
- The impact analysis shows that the PSNP had no impact on consumption poverty based on the national food poverty line and subjective poverty.
- The report also presents descriptive results on agency and intrahousehold decision making.
- PSNP households are less likely to feel they have full control over their life compared to non-PSNP households. And improvement over time has been minimal.
- In terms of speaking up in public – for example, on matters relating to infrastructure, payments in the Public Works, or if some aspects of the PSNP are not implemented fairly or correctly – more than 40 percent of households are not at all comfortable, about 30 percent are comfortable, and the remaining 26–27 percent would speak up with difficulty.

***RQ4: Has PSNP improved human development outcomes?***

*Pregnant women and the PSNP*

- Almost one-third (30 percent) of the women were pregnant after May 2016 and about 10 percent were pregnant at the time of the end-line survey. Among pregnant women, 43 percent were working on Public Works when they learned about the pregnancy. About 60 percent of women also reported that someone from their household (other than themselves) was working on Public Works when they learned about the pregnancy.
- Almost two-thirds (63 percent) of women working on Public Works stopped working when they found out about their pregnancy. Among the women who did not stop working, about 18 percent continued to work until they gave birth.
- On average, women stopped working when they were four months pregnant and about 45 percent wished they could have stopped earlier.
- About 30 percent of women reported receiving direct support benefits after they stopped working. And about 24 percent reported that other household members worked more to compensate for the woman not working.
- About two-thirds (64 percent) reported returning to Public Works after giving birth. On average, women went back seven months after giving birth, and 68 percent of these women said that going back to work affected their ability to breastfeed their child.

*Antenatal care seeking*

- Among those who were pregnant in the last three years, 28 percent were visited by a health worker/community worker and about 50 percent reported receiving antenatal care during that pregnancy.
- About 31 percent reported being pregnant during the COVID period and about 21 percent reported receiving antenatal care during the pregnancy. We do not observe significant differences between PSNP and non-PSNP households.

*Women's knowledge of infant and young child feeding*

- Women's knowledge of breastfeeding is better than their knowledge of timely introduction of complementary foods. The only difference between PSNP and non-PSNP mothers is in breastfeeding knowledge, as 65 percent of the former know about immediate initiation of breastfeeding versus 62 percent of the latter. And 61 percent of PSNP mothers know about exclusive breastfeeding until six months versus 58 percent of non-PSNP mothers. Although these differences are statistically significant, the absolute differences are not large.

*Child health*

- Over one-half (57 percent) of children under five years had a health card. However, a much smaller percentage of children were measured. About 16 percent, 13 percent, and 18 percent were measured for weight, height, and mid-upper arm circumference (MUAC), respectively. It appears that children among PSNP households are slightly more likely to be measured compared to those that belong to non-PSNP households.

*Health extension worker and Health Development Army*

- On average, 40 percent of respondents reported knowing the health extension worker (HEW) – PSNP households were more likely than non-PSNP households to know the HEW (42 percent versus 38 percent, $p<0.01$). Among those who know the HEW, about 30 percent were visited by the HEW at home in the past three months – PSNP households were more likely to report this than non-PSNP households (31 percent versus 26 percent, $p<0.05$).
- During the last home visit, 63 percent reported that the HEW talked about breastfeeding, child feeding, and nutrition. Compared to pre-COVID times, 64 percent of PSNP respondents reported that the frequency of HEW home visits increased.
- About 7 percent of PSNP households are members of the health development army (HDA) and about 16 percent know an HDA member.

*Child labor*

- While examining work performed by children (7–17 years old) on agricultural activities, non-agricultural activities, casual/part-time work, salaried, and PSNP Public Works, we find:
    - About one-quarter of households reported that some child members (7–17 years old) did engage in agricultural activities in the last seven days. Both the fraction of PSNP households

reporting such participation and the average length of participation reported for children in the 7–14 years age group are lower compared to non-PSNP households. The difference, though not large, is statistically significant.

- Some PSNP households report that their children in both the 7–14 and 15–17 age groups did participate in Public Works. Nevertheless, both the fraction of households (4 percent for 7–14-year-olds and 5.2 percent for 15–17-year-olds) and the average length of participation (respectively, 0.31 hours and 0.49 hours in the last seven days) are small.

*Mid-upper arm circumference*

- The mean MUAC among children is 13 and 23 among women.
- About 17 percent and 14 percent of children under five years among PSNP households and non-PSNP households, respectively, suffer from moderate acute malnutrition (MAM). (The difference between PSNP and non-PSNP households is statistically significant.)
- About 7.5 percent children under five years suffer from severe acute malnutrition (SAM).
- The impact estimates do not show an impact of the PSNP on MUAC among children or women. We also do not observe an impact from the PSNP on the prevalence of MAM or SAM among children.

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

- The report presents descriptive and impact results on labor allocation and work intensity in the seven days prior to the end-line survey. The activities included in this section 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-third (30 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.
- 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 one out of five activities in the last seven days. This is slightly higher among PSNP households (1.2) than among non-PSNP households (0.9), with some variation across regions and sex of the household head.
- On average, 61 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 17 percent of households; as expected, this is concentrated among PSNP households (36 percent; refer to Table 6.1).

- The impact analysis shows that participation in the PSNP had a positive and statistically significant impact on total hours spent in the last seven days by the household.
- PSNP also had a positive and statistically significant impact on the total number of activities in which the household was engaged in the previous seven days – 0.4 activities, or about 45 percent of the comparison mean. A large positive impact also exists on the likelihood of being engaged in PSNP Public Works in the last seven days.
- However, there is evidence of reallocation of time across activities among those who were engaged in some labor activity in the last seven days. We find that the PSNP had a negative impact on the proportion of hours spent on agricultural and nonagricultural activities and a positive impact on the proportion of time spent on Public Works.

***RQ8: How are impacts differentiated by household characteristics?***

Throughout the report we present descriptive results disaggregated by gender of the household head and by region. Some of the key results follow:

- While female-headed households report a higher mean food gap, regardless of their PSNP status, the mean food gap among female-headed PSNP households declined by 0.3 months between baseline and end-line but remained unchanged among non-PSNP female-headed households.
- Regional disaggregation suggests the mean food gap increased slightly for non-PSNP households in all regions. In contrast, it declined for PSNP households in Oromia and SNNP, but slightly increased in Amhara by about nine days per year.
- Female-headed households in the PSNP sample had relatively lower mean HDDS than female-headed households in the non-PSNP sample. However, the improvement in mean HDDS is slightly higher for PSNP than for non-PSNP female-headed households.
- Important regional differences exist in mean per capita monthly food, nonfood, and total consumption expenditure levels and growth rates. Overall, households in Amhara have a higher level of mean per capita monthly food expenditures. The growth rate in mean per capita monthly food expenditures for all households is also higher in Amhara, followed by SNNP. Oromia has the lowest growth rate in mean per capita food expenditure. While Amhara's growth rate in mean per capita nonfood expenditures was positive, Oromia and SNNP experienced declines in this indicator.
- Female-headed households are less endowed with livestock assets than their male counterparts, and female-headed PSNP households are even less endowed with livestock compared to non-PSNP female-headed households.
- Male-headed households are better off than female-headed households from the perspective of subjective poverty, but over time both groups witnessed some improvement. The three

regions had very similar levels of subjective poverty at baseline. Over time, Amhara and SNNP saw some improvement.

- 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 had seen improvement.
- In terms of "say in decision making" – on matters relating to agriculture, livestock, employment, and household expenditure – overall, male heads are more likely to be able to make their own decisions than are spouses of both male and female heads.
- In addition, the report presents impact estimates examining whether those households that were worse off at baseline (in terms of food gap and livestock holdings) experienced a differential impact of the PSNP compared to those households that were better off at baseline. This was done by dividing the sample using baseline median values of the food gap and TLU and estimating the PSNP's impact in each of these subsamples on a limited set of outcomes.
- Households that had a food gap greater than the median food gap at baseline (that is, households that were worse off at baseline) experienced no impact from the PSNP on consumption expenditures. In contrast, those households that had a food gap lower than the median food gap at baseline experienced a negative impact on consumption expenditures.
- It also appears that the positive impact on TLU holdings was concentrated among households whose food gap at baseline was lower than the median food gap.
- Households that had TLU holdings less than the median at baseline (that is, households that had fewer livestock at baseline) experienced a decline in the food gap as a result of their participation in the PSNP. Those households that had TLU holdings higher than the median at baseline experienced no impact on their food gap.
- Those with low levels of TLU holdings experienced no impact on consumption expenditures, while households with TLU holdings greater than the baseline median had a negative impact on consumption expenditures.
- We also observe a positive impact on income diversification among households that had TLU holdings lower than the baseline median.

# 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? |
| RQ4 | Has PSNP improved human development outcomes? |
| RQ8 | How are impacts differentiated by household characteristics? |
| 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 8$^{th}$ 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. 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. This evaluation is based on quantitativeand qualitative data collected in three rounds: a baseline survey in 2016, a midline survey in 2018 andan 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, after the program was in operation for five years.

The quantitative end-line survey was originally scheduled for 2020 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 in April-May. This is close to the timing of the 2018 midline survey which was May-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 Highlands was conducted through in-person interviews in Amhara, Oromia, SNNP and Sidama.[1] For security reasons, the 2021 quantitative survey was not conducted in Tigray. Data collection took place from 13 April to 15 May, 2021. In a few areas, the survey was completed several days later due to delays caused by logistical and security issues. 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 Working Group (DWG). 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 public works (PW) and direct support (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.

[1] Sidama Region was formed in June 2020 from SNNP Region. For the purpose of this report, we will refer to the regions that existed in 2016, when the study begun and the sample was drawn.

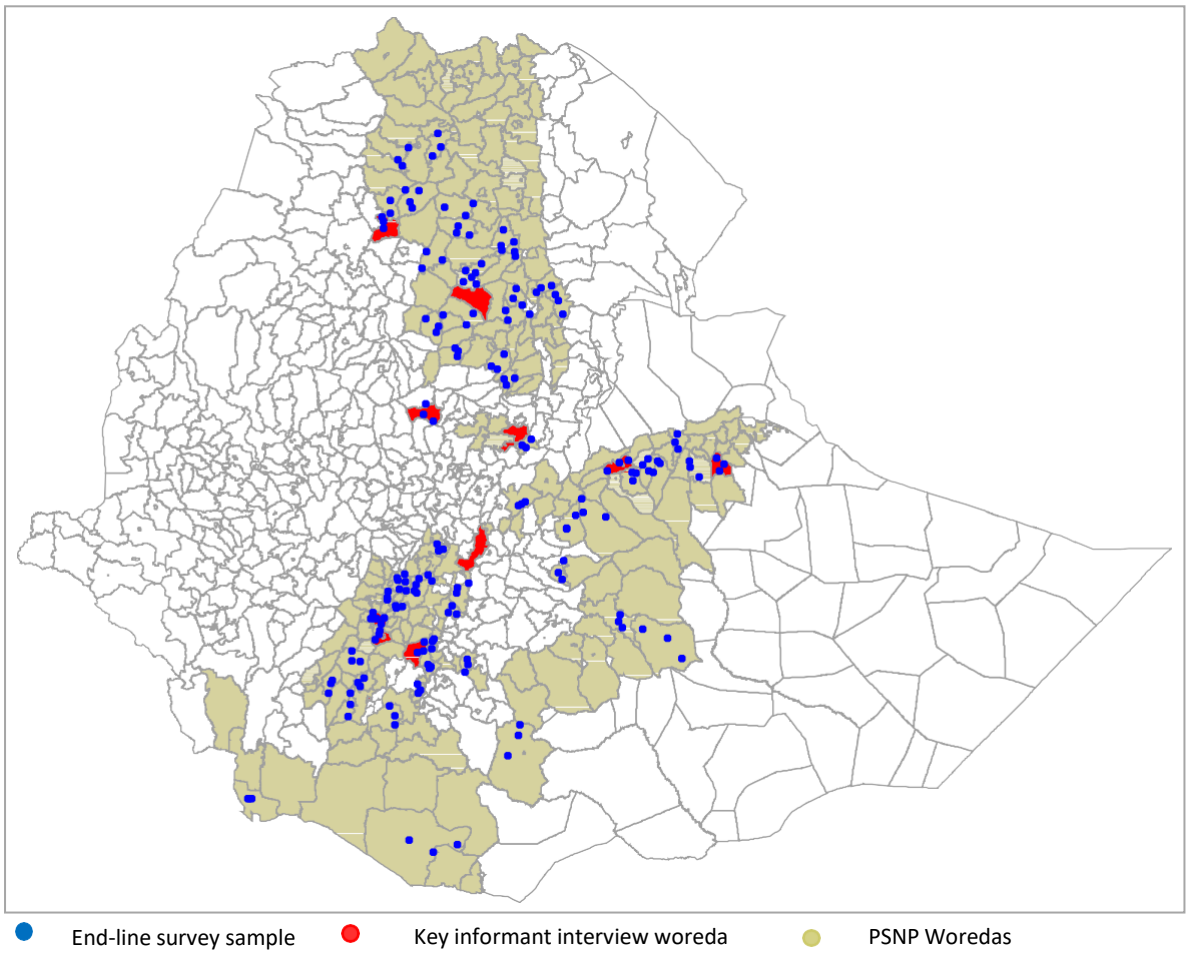
## 2.2 Sampling

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 88 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 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 bottom four rungs of the subjective wealth ranking in the same woredas.

In 2016, the total number of households interviewed during the baseline survey was 7,291 in Amhara, Oromia, SNNP and Tigray. These households were sampled from 264 kebeles or enumeration areas in 88 woredas. In 2018, we re-interviewed 6,998 households in the Highland households. The 293 households not interviewed represent an attrition rate of 4.0 percent (or two percent per year). Attrition was highest in Oromia (5.5%) and lowest in SNNP (2.9%) with some of the attrition due to security problems that prevented CSA teams from visiting survey sites in Darolebu woreda in Oromia region and Yirgacheffe woreda in Amhara. The remaining attrition was approximately random. The target sample for the 2021 survey in the highlands included all households in the baseline highlands sample. Figure 2.1 shows the location of the quantitative survey enumeration areas in the highlands for the end-line survey, with the Tigray sample excluded.

**Figure 2.1: Enumeration areas in the quantitative survey sample and qualitative survey woredas - highlands**




## 2.3 The 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. Cash payments (including payment schedules, attendance sheets, obtaining funds, making payments)

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

Interviews were completed by experienced survey supervisors who were instructed to meet with staff of 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 livelihoods transfer 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, a 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 had to include at least one member of the kebele Food Security Task Force and at least one woman and they were 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 done in previous years, the community questionnaire covered the following topics: location and access; water and electricity; services; education and health facilities; production and marketing; migration; 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; graduation; 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 7,291 households in Amhara, Oromia, SNNP and Tigray. These households were sampled from 264 kebeles or enumeration areas in 88 woredas. In the 2021 end-line survey, it was not possible to interview households in Tigray, removing 1,794 households from the sample. In addition, 35 households from the baseline had incomplete interviews or refused to consent to be interviewed. Thus, the target end-line sample was 5,462 households from baseline in Amhara, Oromia and SNNP. Table 2.1 shows the number of baseline households reached in the end-line survey by region. Enumerators were able to interview 5,082 households in the end-line survey. Attrition was highest in Oromia at 10.9 percent and lowest in SNNP at 3.0 percent. Overall, the attrition rate was 7.0 percent. This is a relatively low attrition rate over a 5-year period. Nonetheless, in Appendix B, we assess whether households that were not interviewed in 2021 are systematically different to the households that were interviewed.

We find that female headed households are more likely to attrit from the sample and that the probability of attriting is declining in household head age. Somewhat surprisingly, households with

more land are significantly more likely to attrit. Households with higher quality housing are less likely to attrit than others.[2] Also, households that experienced a non-drought shock in the last two years are more likely to attrit, but those living in a kebele that received humanitarian assistance in response to a drought in 2015 are less likely to have dropped from the sample, presumably because they are less likely to have moved. Finally, households in Oromia and SNNP are more likely to attrit than those in Amhara.

**Table 2.1: End-line survey household sample**

| | Amhara | Oromia | SNNP | Total |
|---|---|---|---|---|
| Interviewed | 1,704 | 1,591 | 1,787 | 5,082 |
| Attrited | 130 | 195 | 55 | 380 |
| Attrition level (%) | 7.1 | 10.9 | 3.0 | 7.0 |
| Total | | | | 5,462 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

The household questionnaire is a multi-topic instrument that collected 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 highlands.

[2] Housing quality was assess using two variables: (i) the enumerator's observation of whether the housing was in moderate to very good condition, and (ii) whether the household had a corrugated or tine roof.

**Table 2.2: Contents of the household questionnaire fielded in the Highlands**

| Module | Title | Description | Changes since baseline (2016) |
|---|---|---|---|
| **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. | No major changes |
| 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. | No major changes |
| 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. | No major changes |
| 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. | No major changes |
| 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. | This module has the largest number of changes. First, payments for PW and DS are now combined (consistent with what appears on the PSNP 4 client card.) Additional questions are asked to make sure we are covering points listed in the TOR which were not covered in 2016. |

| Module | Title | Description | Changes since baseline (2016) |
|---|---|---|---|
| | | | Second, we know that from the 2016 survey round that targeting of the PSNP was pretty good in the Highlands. Because of this, and because we want to add new questions about payment processes and graduation, many questions about targeting processes were dropped.<br>Third, new questions on experiences with payments, especially with respect to e-payments. Fourth, questions about graduation added. |
| 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. | No major changes |
| 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. | No major changes |
| **Adult female questionnaire** | | | |
| F1 | Time use, employment, own business activities and credit and savings | This module collects information on women's time use, employment in the last 12 months, own business activities, and credit and savings for the main adult female. | No major changes |
| F2 | Homestead gardens | This module covers homestead garden production. | No major changes |
| F3 | Assets, livestock, livestock production, livelihoods component | This module collects data on assets owned by the spouse and her involvement in livestock production and activities under the livelihood component. | No major changes |
| F4 | Housing and water | This module covers WASH (water, sanitation and hygiene). | No major changes |

| Module | Title | Description | Changes since baseline (2016) |
|---|---|---|---|
| F5 | Non-food expenditures | This short module collects data on household expenditures on household consumables that were not covered in module M6. | No major changes |
| F6 | Food consumption | This module covers the food consumption by the main female the day before the interview. It also collects data on household's food consumption in the past 7 days. The female respondent is also asked to respond to questions regarding food security in the last 12 months. | No major changes |
| F7 | Decision-making and voice at home and in the community | This short module asks about main females' views about decision making and voice at home and in the community. | Questions about Locus of Control were dropped for the midline survey round but were included again in this end-line survey round (2021). |
| F8 | Health and nutrition | This module has 6 sections:<br>PSNP during pregnancy and lactation (up to one year after childbirth); Use of antenatal and postnatal services; Infant and young child feeding (IYCF) practices; Child health history; Maternal IYCF knowledge and perceptions; Exposure to health and nutrition services. | After the 2016-baseline, IFPRI was tasked by the Bill and Melinda Gates Foundation (BMGF) to conduct a separate quantitative impact evaluation of the nutrition sensitive components of the PSNP 4. Two baseline surveys were conducted in 2017, one in March and another one in August. These surveys took place in the same 264 highland enumeration areas as the main PSNP 4 evaluation surveys. The end-line surveys of the BMGF funded study were completed in 2019. Given the extent of this BMGF funded nutrition study, a reduced version of the nutrition module was fielded in this midline survey. Consequently, for example, child and maternal anthropometric (height, weight) data were not collected. |

## 2.4 Impact evaluation strategy

### *2.4.1 Overview*

As is now well-understood, 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 – as is the case with the PSNP - 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 from data on non-beneficiaries who are observationally similar to the beneficiary group.

As we did in the 2018 Midline Outcomes report, the approach we use to estimate impact combines difference-in-difference and matching methods. Matching approaches are appropriate for programs like the PSNP because targeting of beneficiaries at the community and household level is conducted by the program (and thus cannot be subjected to randomization) and because a regression discontinuity design is infeasible (the use of community targeting means that there is no unique eligibility cut-off separating beneficiaries from non-beneficiaries). Matching methods construct the counterfactual by matching program beneficiaries to non-beneficiaries using observed characteristics; program impacts are estimated 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. Heckman, Ichimura, and Todd (1998), Heckman, Ichimura, and Todd (1997) 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 andworedas.

We end by noting several key assumptions. First, following the principles of 'one-PSNP', we estimate the average impact of the whole program; Public Works (PW) and Direct Support (DS) combined. Second, these matching models require the basic assumption, known as unconfoundedness, that, after controlling for observables, mean outcomes for non-beneficiaries are identical to outcomes of beneficiaries if they had not received the program. Whether this assumption holds depends on whether observable variables are primarily responsible for determining participation in the program and related outcomes. This is a plausible assumption for the PSNP data collected using the same survey questionnaire in this context, in which relatively poor households were screened before sampling as non-beneficiaries for the study. Finally, other development and humanitarian programs

operate in rural Ethiopia in parallel to the PSNP. As a result, some of our control households may have directly or indirectly benefitted from non-PSNP related activities taking place in their localities. Since these programs are not part of the PSNP, we consider this as a normal operating environment. In this sense, the comparison group does not represent a 'pure' control group with no other programs, but rather a comparison group that captures 'business as usual.' Thus, our impact estimates represent the impact of PSNP4 over and above of any other activities taking place in the PSNP areas.

### *2.4.2 Definition of treatment status*

An important feature of the evaluation is the definition of treatment status. PSNP4 has been operating for 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 end-line survey has several variables that are useful in defining beneficiary status. First, households are asked if they participated in Public Works (PW) in 2021 and in each of the previous three years.[3] Households also reported if they received payments from Public Works for each of these 4 years. Households were also asked to report if they received payments from the Direct Support (DS) component of the program, which does not include a work requirement.

We used these data to calculate an indicator of the number of years that each household had participated in the PSNP4 out of the last four years, where a household was considered to have participated in any given year if they indicated taking part in PW or receiving any payment from DS in that year. We also constructed a PSNP4 annual payments indicator variable which showed if a household had received any PW or DS payments during that year.

Table 2.3 summarizes these variables. In Panel A, we summarize the number of years in which a household participated in PW or DS against the variable for whether that household received any PSNP4 payment in 2021. There we see that the vast majority of households who participated in PSNP4 did so for all four years (n=2,057). Only 348 households took part in PSNP4 but participated for fewer than four years. Also, 2,706 households did not participate in PSNP4 in any year; these are the comparison group households. This shows that beneficiary status has been remarkably stable. Panel A suggests that using the requirement that a household participated in PSNP4 for 3 or 4 years could work well in capturing the majority of participants who stayed in the program for the entire period. Panel A also shows that we add a requirement that the household must have been paid by PSNP4 in the current year to be considered for the beneficiary group for this evaluation, the sample will lose 269 households who participated all four years but did not get paid in the first half of 2021.

Panel B of Table 2.3 further shows that if we define beneficiaries those that participated in PSNP4 for 3 or 4 years and also add the eligibility criterion that they received payments in 2020 or 2021, very little is gained by adding the payment requirement because this would add only 57 households to the

[3] Respondents were asked about PSNP participation for the last four years. Earlier years were omitted out of concern that errors in recall would be high.

beneficiary group. Thus, we define beneficiary status as participating in PSNP4 for either 3 or 4 years and comparison group status as not participating or receiving PSNP4 payments in any of those years.[4] It is helpful to note that more than half of those households that participated for three years were participating in 2021, so those households likely joined the program a bit late. A strength of this definition is that it is simple, and it captures meaningful participation. In the Midline Outcomes Report analysis, we defined beneficiary status as including both participation in PW and also having at least one other measure indicating involvement, including payments or indicating that you had first been selected by your kebele leaders. However, with participation relatively stable in the period since the midline survey and payments less stable over that period, a beneficiary definition at end-line based on participation alone is effective.

**Table 2.3: Years of PSNP4 participation and payments**

| | Number of years the household participated in PSNP, either PW or DS | | | | | |
|---|---|---|---|---|---|---|
| | **0** | **1** | **2** | **3** | **4** | **Total** |
| Panel A<br>Received PSNP4 payments in 2021 | | | | | | |
| No | 2,706 | 28 | 37 | 59 | 269 | 3,099 |
| Yes | 0 | 52 | 39 | 133 | 1,788 | 2,012 |
| Total | 2,706 | 80 | 76 | 192 | 2,057 | 5,111 |
| | | | | | | |
| Panel B<br>Received PSNP4 payments in 2020-2021 | | | | | | |
| No | 2,706 | 17 | 15 | 10 | 47 | 2,795 |
| Yes | 0 | 63 | 61 | 182 | 2,010 | 2,316 |
| Total | 2,706 | 80 | 76 | 192 | 2,057 | 5,111 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

### *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.

[4] We chose to represent active participation in the PSNP as participation for 3 or 4 of the last 4 years, accepting any pattern over those years, believing that this represents substantial participation regardless of which year was missed.

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.[5]

### *2.4.4 Baseline household and community variables used in matching*

The baseline households and community variables used in the matching models are summarized in Table 2.4. These variables include measures of household head characteristics, household demographics, assets, housing characteristics, exposure to shocks, kebele characteristics, and regional dummy variables.

Table 2.4 reports two measures of statistical significance for the hypothesis test that the mean baseline characteristics are equal between PSNP and non-PSNP households. P-values are from a t-test of equality of means. A p-value below 0.05 indicates a significant difference in means at the 5 percent level. The normalized difference for each outcome in each sample is defined as the difference in sample means between the PSNP and non-PSNP samples divided by the average standard deviation in PSNP and non-PSNP samples (Imbens 2015). The normalized difference is a measure of significance that is not affected by sample size. As suggested by Imbens and Rubin (2015), in Table 2.4 we interpret normalized differences of 0.25 or above as indicating a significant difference in means between the PSNP and non-PSNP samples.

Table 2.4 shows that PSNP households are more likely to be female headed than non-PNSP households. Household heads in the PSNP are slightly older as well. The demographic composition of PSNP and non-PSNP households is similar in terms of gender and number of adults and dependents. Land holdings are smaller for PSNP households than non-PSNP households and an index of productive assets constructed by principal components analysis (PCA) is also lower for PSNP households. PSNP households are less likely to live in better-quality housing or have improved roof materials. Exposure to non-drought shocks is similar across the two samples. Non-PSNP households live in kebeles with significantly better-quality roads on average. PSNP households live in kebeles that are more likely to have received temporary drought or humanitarian relief in 2015. Although PSNP and non-PSNP households are sampled from the same villages, this pattern of differences in community level variables for PSNP and non-PSNP households can arise because in practice PSNP and non-PSNP households are not always sampled in fixed proportions.

[5] Assuring balance of the matching covariates within intervals of the predicted propensity score distribution ensures that treated households at each interval of the predicated propensity score has comparison group households with similar characteristics and similar predicted probability of being in the program.

**Table 2.4: Summary statistics of baseline covariates of PSNP and non-PSNP samples**

| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Female headed household | 0.292 | 0.455 | 2,235 | 0.208 | 0.406 | 2,691 | 0.000 | 0.193 |
| Household head's age | 46.243 | 14.970 | 2,235 | 44.639 | 15.152 | 2,691 | 0.000 | 0.107 |
| Household head's age squared | 2,362.454 | 1,572.804 | 2,235 | 2,222.156 | 1,553.187 | 2,691 | 0.002 | 0.090 |
| Number of males | 2.347 | 1.468 | 2,235 | 2.489 | 1.478 | 2,691 | 0.001 | -0.097 |
| Number of females | 2.498 | 1.363 | 2,235 | 2.473 | 1.334 | 2,691 | 0.532 | 0.018 |
| Number of members 16 to 60 | 2.183 | 1.122 | 2,235 | 2.289 | 1.121 | 2,691 | 0.001 | -0.093 |
| Number of dependents | 2.661 | 1.712 | 2,235 | 2.674 | 1.684 | 2,691 | 0.785 | -0.008 |
| Total land holdings in hectare | 0.712 | 0.982 | 2,235 | 0.821 | 0.970 | 2,691 | 0.000 | -0.111 |
| HH productive asset PCA | -0.409 | 2.595 | 2,235 | 0.005 | 2.632 | 2,690 | 0.000 | -0.158 |
| Housing in moderate to very good condition | 0.632 | 0.482 | 2,235 | 0.747 | 0.435 | 2,690 | 0.000 | -0.250 |
| Household has corrugated metal roof | 0.272 | 0.445 | 2,235 | 0.404 | 0.491 | 2,690 | 0.000 | -0.282 |
| Household experienced any non-drought shocks in the last 2 years | 0.408 | 0.491 | 2,235 | 0.417 | 0.493 | 2,691 | 0.491 | -0.020 |
| Community is connected to a road made of stone | 0.461 | 0.499 | 2,235 | 0.488 | 0.500 | 2,691 | 0.059 | -0.054 |
| Community is connected to a dirt road | 0.454 | 0.498 | 2,235 | 0.404 | 0.491 | 2,691 | 0.000 | 0.101 |
| Road is accessible in rainy season | 0.294 | 0.455 | 2,235 | 0.364 | 0.481 | 2,691 | 0.000 | -0.151 |
| Number of Development Agents in the kebele | 3.324 | 1.496 | 2,235 | 3.190 | 1.531 | 2,691 | 0.002 | 0.089 |
| Kebele received temporary drought/humanitarian relief in 2015 | 0.801 | 0.399 | 2,235 | 0.700 | 0.458 | 2,691 | 0.000 | 0.236 |
| Region: Oromia | 0.381 | 0.486 | 2,235 | 0.294 | 0.455 | 2,691 | 0.000 | 0.187 |
| Region: SNNP | 0.286 | 0.452 | 2,235 | 0.335 | 0.472 | 2,691 | 0.000 | -0.107 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

### *2.4.5 Estimating the propensity score*

Table 2.5 presents the results of the propensity score estimation based on a logit model in which the dependent variable is the indicator for whether the household was a PSNP beneficiary in 2021. 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 in Table 2.5.

**Table 2.5: Logit model for PSNP4 participation**

| Variables | (1) |
|---|---|
| Female headed household | 0.313*** |
| | (0.080) |
| Household head's age | 0.027** |
| | (0.011) |
| Household head's age squared | -0.000** |
| | (0.000) |
| Number of males | 0.035 |
| | (0.026) |
| Number of females | 0.089*** |
| | (0.024) |
| Number of members 16 to 60 | -0.038 |
| | (0.035) |
| Total land holdings in hectare | -0.126*** |
| | (0.038) |
| HH productive asset PCA | -0.064*** |
| | (0.014) |
| Housing in moderate to very good condition | -0.287*** |
| | (0.070) |
| Household has corrugated metal roof | -0.508*** |
| | (0.068) |
| Household experienced any non-drought shocks in the last 2 years | -0.031 |
| | (0.061) |
| Community is connected to a road made of stone | -0.091 |
| | (0.071) |
| Community is connected to a dirt road | 0.189*** |
| | (0.073) |
| Road is accessible in rainy season | -0.152** |
| | (0.070) |
| Number of Development Agents in the kebele | 0.019 |
| | (0.024) |
| Kebele received temporary drought/humanitarian relief in 2015 | 0.468*** |
| | (0.071) |
| Region: Oromia | 0.675*** |
| | (0.088) |
| Region: SNNP | 0.169** |
| | (0.080) |
| Constant | -1.448*** |
| | (0.275) |
| Number of observations | 4,925 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Following Dehejia and Wahba (1999, 2002), the balancing routine used to confirm the covariates used in this model tests for equality of 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.

Next, we consider the degree of common support or overlaps in the predicted propensity score distributions between the PSNP and non-PSNP samples. Figure 2.2 shows the propensity scores estimated on the full sample. Figure 2.3 shows the same results but with 5% of the observations trimmed from the common support regions by cutting 2.5% at each tail. In both figures, the propensity score model appears to be well behaved. Beneficiary observations appear to have many non-beneficiary comparison observations at similar PPS for most of the distribution of the PPS.

**Figure 2.2: The distribution of the predicted propensity score by PSNP beneficiary status, full sample**

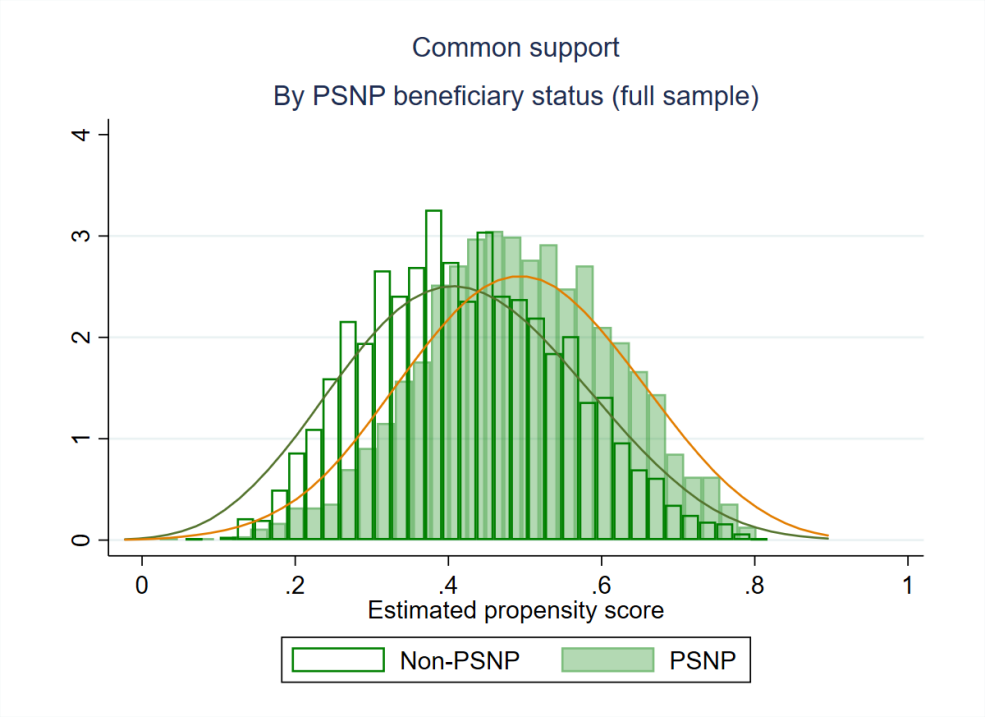
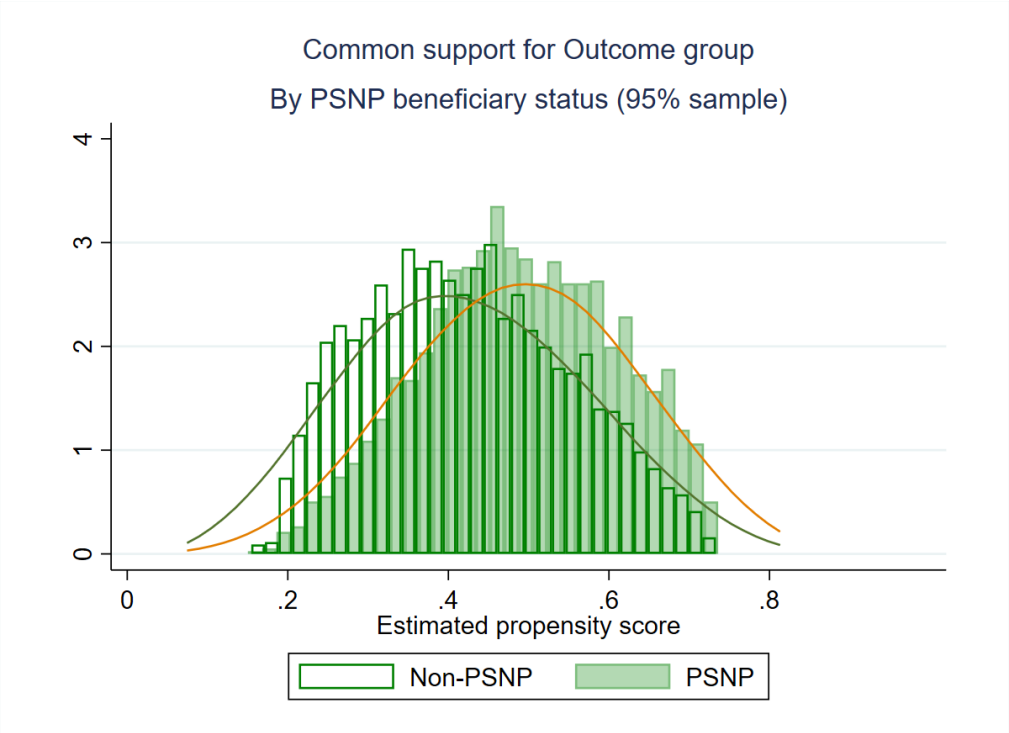


**Figure 2.3: The distribution of the predicted propensity score by PSNP beneficiary status, 95% trimmed sample**




*2.4.6 The choice of matching algorithm*

We generate our impact estimates using nearest neighbor matching (NNM) (Abadie and Imbens (2006)). Like the better-known Propensity Score Matching (PSM) technique, NNM matches each beneficiary household to one or more non-beneficiary households based on pre-program household and community characteristics. However, NNM matches directly on the variables themselves by selecting non-beneficiaries for the match that minimize the average difference in characteristics from the beneficiary using a multidimensional metric to determine the weights for constructing theaverage. The effect of participating in the program is measured as the average difference in the outcome for each beneficiary from the average outcome among its matched non-beneficiaries. Following the approach taken in the 2018 Midline Outcomes Report, we will use five neighbors in ourNNM models and because we match on more than two continuous variables, we use the bias correction method proposed by Abadie and Imbens (2006).

Differences between NNM and PSM derive primarily from the rule used to select comparable non-beneficiaries and the weights used to construct the difference in weighted average outcomes. NNM matches beneficiaries to non-beneficiaries based directly on the observable characteristics. Each beneficiary is matched to the group of non-beneficiaries with the smallest average difference in pre-program characteristics, where this difference is determined using a multi-dimensional metric across all control variables. The advantage of the NNM method is that it permits the calculation of analytical standard errors. For PSM, the standard errors need to be computed using bootstrap methods (Brownstone and Valletta 2001) – an approach that is not recommended in the context of matching (Abadie and Imbens 2008).

Finally, we estimated the impact of PSNP4 by NNM using the teffects nnmatch routine in STATA. After running the NNM model, we conducted a balance check to confirm that the mean of the retained X variables are balanced between the PSNP and non-PSNP samples when the weights from the NNM routine are applied. The results of these balance tests are presented in Table 2.6 below. In this table, the mean, SD and N are reported for the final set of X variables in the common support sample for the PSNP sample and the non-PSNP sample. The normalized difference is reported for the test of equality of means on the raw data shown in the table and then again using the matched data with weights applied. Most of the normalized differences in the matched data are very small and none of them is larger in absolute value than 0.25, suggesting that these models are balanced.

**Table 2.6: Balance table for baseline covariates in Food Security outcome, by PSNP and non-PSNP samples**

| | PSNP households | | | Non-PSNP households | | | Normalized difference (raw) | Normalized difference (matched) |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Female headed household | 0.292 | 0.455 | 2,235 | 0.208 | 0.406 | 2,689 | 0.194 | 0.042 |
| Household head's age | 46.243 | 14.970 | 2,235 | 44.634 | 15.155 | 2,689 | 0.107 | 0.025 |
| Household head's age squared | 2,362.454 | 1,572.804 | 2,235 | 2,221.750 | 1,553.528 | 2,689 | 0.090 | 0.015 |
| Number of males | 2.347 | 1.468 | 2,235 | 2.490 | 1.478 | 2,689 | -0.097 | -0.018 |
| Number of females | 2.498 | 1.363 | 2,235 | 2.473 | 1.334 | 2,689 | 0.017 | 0.022 |
| Number of members 16 to 60 | 2.183 | 1.122 | 2,235 | 2.289 | 1.121 | 2,689 | -0.093 | -0.007 |
| Number of dependents | 2.661 | 1.712 | 2,235 | 2.674 | 1.684 | 2,689 | -0.111 | -0.078 |
| Total land holdings in hectare | 0.712 | 0.982 | 2,235 | 0.821 | 0.970 | 2,689 | -0.159 | -0.063 |
| HH productive asset PCA | -0.409 | 2.595 | 2,235 | 0.007 | 2.632 | 2,689 | -0.252 | -0.046 |
| Housing in moderate to very good condition | 0.632 | 0.482 | 2,235 | 0.747 | 0.435 | 2,689 | -0.282 | -0.075 |
| Household has corrugated metal roof | 0.272 | 0.445 | 2,235 | 0.404 | 0.491 | 2,689 | -0.020 | -0.013 |
| Household experienced any non-drought shocks in the last 2 years | 0.408 | 0.491 | 2,235 | 0.418 | 0.493 | 2,689 | -0.054 | 0.003 |
| Community is connected to a road made of stone | 0.461 | 0.499 | 2,235 | 0.488 | 0.500 | 2,689 | 0.101 | 0.016 |
| Community is connected to a dirt road | 0.454 | 0.498 | 2,235 | 0.404 | 0.491 | 2,689 | -0.151 | -0.026 |
| Road is accessible in rainy season | 0.294 | 0.455 | 2,235 | 0.364 | 0.481 | 2,689 | 0.089 | -0.009 |
| Number of Development Agents in the kebele | 3.324 | 1.496 | 2,235 | 3.190 | 1.532 | 2,689 | 0.236 | 0.079 |
| Kebele received temporary drought/humanitarian relief in 2015 | 0.801 | 0.399 | 2,235 | 0.700 | 0.459 | 2,689 | 0.188 | 0.035 |
| Region: Oromia | 0.381 | 0.486 | 2,235 | 0.293 | 0.455 | 2,689 | -0.108 | -0.022 |
| Region: SNNP | 0.286 | 0.452 | 2,235 | 0.335 | 0.472 | 2,689 | | |

***Notes***: Estimates from the comparison and beneficiary groups selected by the matching model.

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

### *2.4.7 Robustness checks*

We ran a number of robustness checks to determine whether the impact estimates are sensitive to the specifications. These robustness checks included: (i) estimating the model on the 95% trimmed common support sample, (ii) modifying the covariate selection, and (iii) estimating all models by another matching method known as inverse probability weighted regression (IPWR). The results presented in this report are fairly robust to changes in the trimming and in the choice of matching estimator. Some of the impact estimates are a bit sensitive to the choice of included baseline covariates. In these cases, we tried multiple specifications to find one that held up under modest changes in specification.

We also considered whether past receipt of the PSNP, during Phase 3, may be biasing the impact estimates if current PSNP4 beneficiaries were more likely to receive PSNP3 and may therefore have a more muted response to PSN4 transfers. Table 2.7 shows the share of PSNP4 beneficiaries in PW and in DS that received transfers under PSNP3, for each number of years from 0-5, by region. The majority of PSNP4 beneficiaries of PW or DS were also beneficiaries of PSNP3, with 17-28% receiving PSNP3 PW transfers for just one year and many more (28-59%) receiving PSNP3 PW transfers for all five years. Thus, the level of household wellbeing at baseline in the PSNP4 evaluation sample is certainly affectedby prior participation in PSNP3. However, the matching methods used control for baseline levels of wellbeing, including landholdings, an index for the value of productive assets and housing characteristics. Matching on these variables helps to capture the effects of any improvements in wellbeing that these households derived from PSNP3, and these beneficiary households should be matched with non-beneficiary households with similar levels of wellbeing. In this sense, the matchingapproach should appropriately capture the beneficial effects of PSNP3 among PSNP4 sample households and should not induce any bias in estimated impacts of PSNP4, unless participation in PSNP3 led to improvements in household wellbeing, such as the food gap or assets, that are not captured by these baseline control variables. We expect that the potential for such bias to be small.

Also, we note that controlling for past participation in PSNP3 in the matching model for impacts of PSNP4 would not be effective because PSNP3 participation is so highly predictive of participating in PSNP4. Including such a control variable would mean screening out a lot of comparison group households and weakening the area of ‘common support,’ which is the region of overlap in predicted probability of participation in PSNP4 between beneficiaries and non-beneficiaries that is needed for effective matching estimates.

**Table 2.7: Composition of PSNP4 PW and DS clients by participation intensities in PSNP3**

| **Public works (PW)** | | | | | | |
|---|---|---|---|---|---|---|
| **What proportion of PSNP4 PW clients have participated 0, 1, … 5, years in PSNP3?** | **0** | **1** | **2** | **3** | **4** | **5** |
| Tigray | 42.0 | 18.0 | 1.0 | 0.0 | 1.0 | 39.0 |
| Amhara | 19.0 | 17.0 | 2.0 | 1.0 | 2.0 | 59.0 |
| Oromia | 36.0 | 21.0 | 4.0 | 4.0 | 3.0 | 33.0 |
| SNNP | 36.0 | 28.0 | 2.0 | 3.0 | 3.0 | 28.0 |
| **Direct support (DS)** | | | | | | |
| **What proportion of PSNP4 DS clients have participated 0, 1, … 5, years in PSNP3?** | **0** | **1** | **2** | **3** | **4** | **5** |
| Tigray | 29.0 | 10.0 | 3.0 | 1.0 | 1.0 | 57.0 |
| Amhara | 26.0 | 7.0 | 2.0 | 2.0 | 3.0 | 59.0 |
| Oromia | 43.0 | 16.0 | 3.0 | 1.0 | 1.0 | 36.0 |
| SNNP | 30.0 | 19.0 | 1.0 | 3.0 | 4.0 | 42.0 |

***Source***: Reproduced from Table 2.3 of the baseline report.

In addition, we considered how impact estimates would be affected if households in the control group participated in other activities or programs. There are several development and humanitarian programs operating in rural Ethiopia in parallel to the PSNP. As a result, some of our controlhouseholds may have directly or indirectly benefitted from non-PSNP related activities taking place in their localities. Similarly, some PSNP beneficiary households also benefit from these other programs.Since these programs are not part of the PSNP, we consider this as a normal operating environment. In this sense, the comparison group does not represent a 'pure' control group with no other programs,but rather a comparison group that captures 'business as usual.' We do not know whether these control group households participate in these other programs at a higher rate, but we control for thebenefit of past participation in such programs by matching on baseline measures of wellbeingincluding assets and housing characteristics. Overall, our impact estimates represent the impact of PSNP4 over and above the effect of any other activities taking place in the PSNP areas.

## Chapter 2 references


Imbens, G. W. 2015. Matching Methods in Practice: Three Examples. *Journal of Human Resources* 50(2):373-419.

Imbens, G. W., and Rubin, D. B. 2015. Causal inference: For statistics, social, and biomedical sciences an introduction, in *Causal Inference: For Statistics, Social, and Biomedical Sciences: an Introduction*.


## 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.[6] 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.

[6] 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 used the cash and grains transfers they received last, which could be a month ago or any time earlier. The chapter will address **RQ17: Use of transfers**. Section 3.2 discusses the proportion of households that used their cash transfers on different items (expenditure categories) and the importance of these items in total cash transfers. For cash transfers, ten expenditure categories are identified in the survey instrument. In the description below we aggregate these into six expenditure categories. Expenditures on non-food items that directly benefit children, male adults, and female adults are aggregated into one category, non-food expenditure. Cash transfers that recipients were asked to give to other households by anyone in a position of authority, which involves only 0.7 percent (14) households, is combined with cash transfers used to help other households into a category labelled as "Other households". Finally, cash transfers given to local authorities, which involves 2.8 percent (57) households, is aggregated with cash transfers given to persons designated to collect PSNP payments into a category labelled as "Others".

Section 3.3 describes how households used the last food transfers they received, which could be a month ago or any time earlier, into one or more use categories. We also describe the importance of these categories in total grains 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 for cash transfers in the last paragraph.

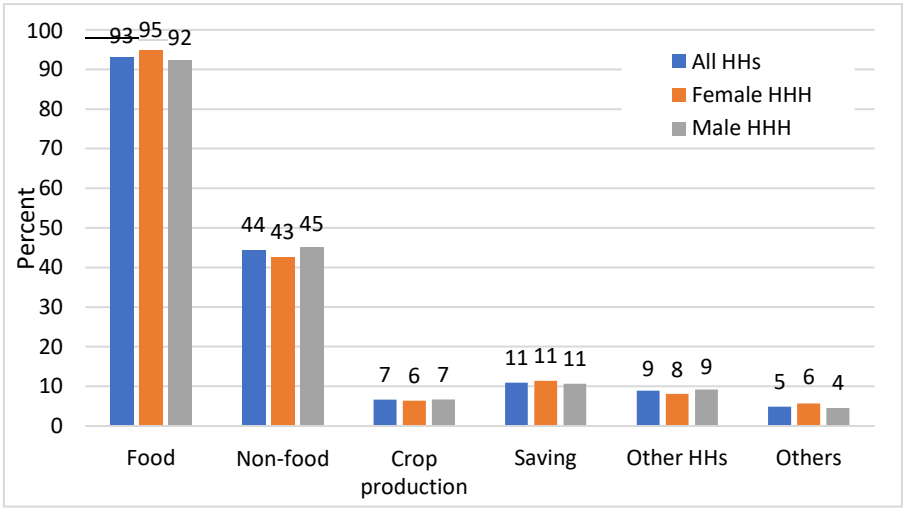
## 3.2 Cash transfers

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. For instance, it describes the percentage of the households that 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 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 use cash transfers for different purposes across gendered household types. Over 93 percent of the households used their cash transfers to buy food. Over 40 percent used cash transfers to buy non-food items, 6 percent used them to purchase inputs used in crop production, 11 saved cash transfers, 9 percent shared them with other households; and 5 percent used cash transfers for other purposes. There are no statistically significant differences in the proportion of gendered household types that used their cash transfer for different purposes with

one exception. The exception is that the proportion of female headed households that purchased food is slightly but statistically significantly higher than the proportion of male headed households.

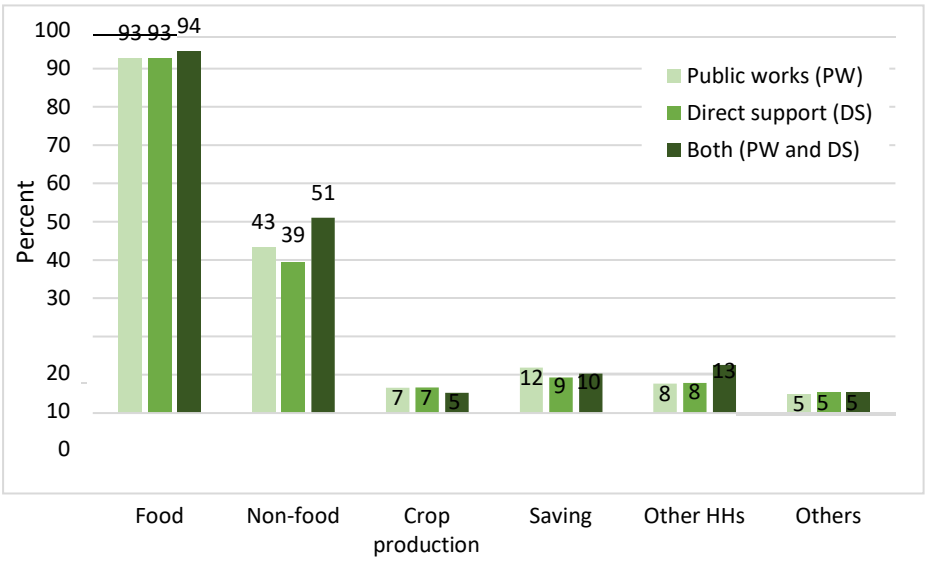
**Figure 3.1: Households that used cash transfer for different purposes (%), by gender of the head of household**




***Note***: HHs = households; HHH = Headed household.
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Figure 3.2 shows the proportion of households that used cash transfers for different purposes across PSNP status.[7] The figure indicates that a slightly higher proportion of households that participate in both public works and direct support use cash transfers to purchase non-food items and to help other households. Both of these differences are statistically significantly different from zero.

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

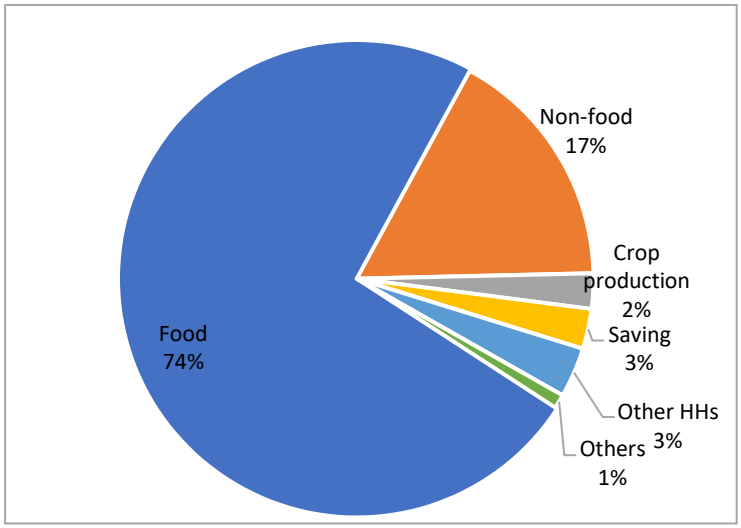



***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

[7] 58.9 percent of the households were in public works, 19.4 percent in direct support, and 21.8 in both direct support and public works.

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 10 percent was spent on non-food goods that directly benefit children while the remaining was about equally shared between non-food items that directly benefit male and female adults. Out of the remaining cash transfers 3.4 percent is used to help other households, 2.7 percent saved, and 2.4 percent used in crop production.

**Figure 3.3: Percentage of cash transfer used for different purposes,**

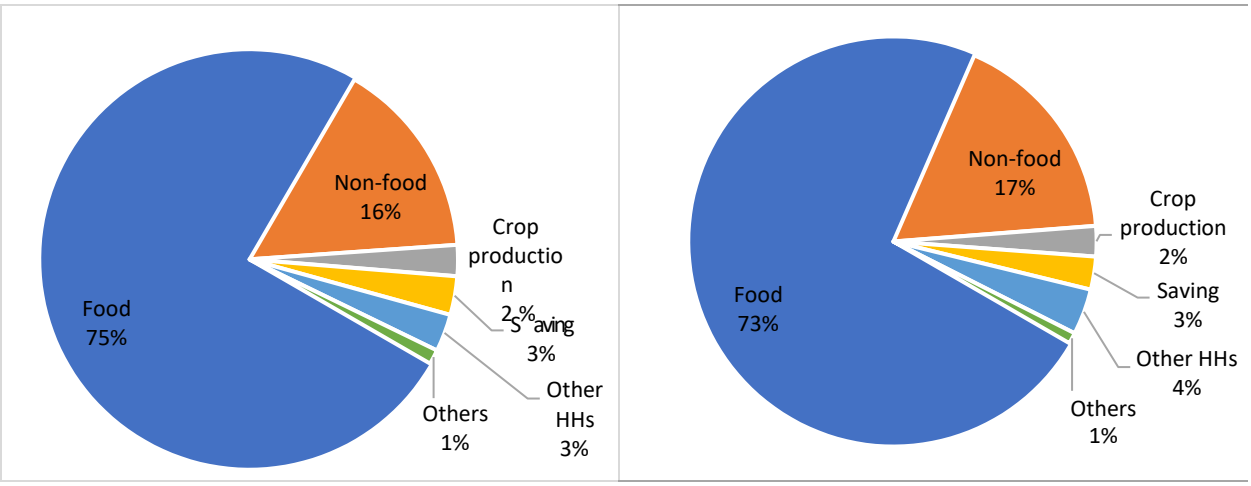



***Note***: HHs = households.
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Figures 3.4 below, which shows the importance of different expenditure categories across gendered household types, is mostly consistent with the implication of Figure 3.1. 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 (Figure 3.4). However, differences in the importance of expenditure categories are not statistically significant across gendered household types. For the most part, Figure 3.5 is consistent with what was observed in Figure 3.2. Relative to households that participate only in public works or in direct support those that participate in both use a higher share of cash transfers to purchase non-food items and save a lower share of their cash transfers and these differences are statistically significant.

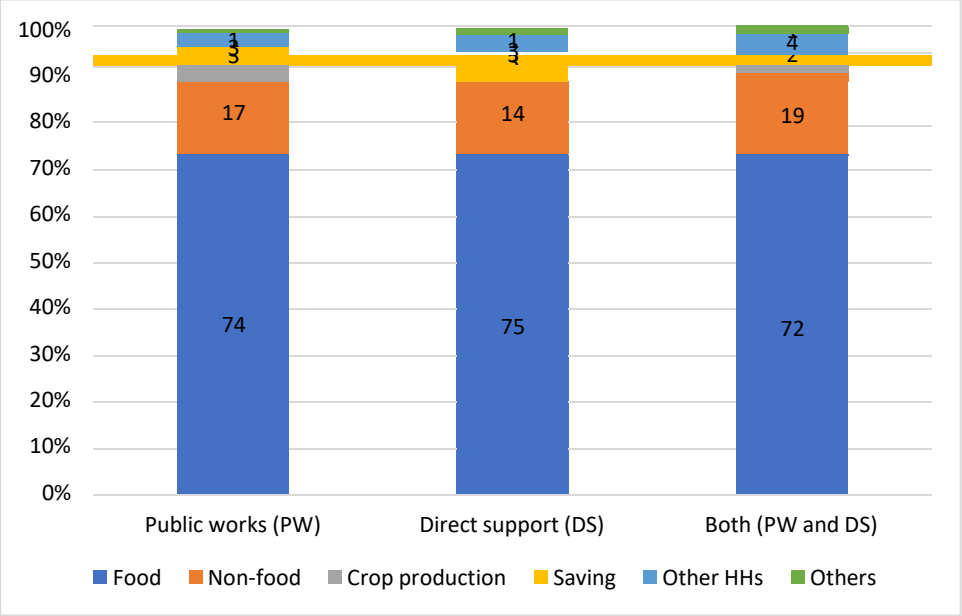
**Figure 3.4: Percentage of cash transfer used for different purposes, by gendered household type**

Female headed households Male headed households




***Note***: HHs = households.
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

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




***Note***: HHs = households.
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

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 results reveal that over 90 percent of households spent cash transfers to buy food and food purchases accounted for 69-77.5 percent of cash transfers. About 40-52.5 percent of the households used cash transfers to buy non-food items and non-food items accounted for 15-19 percent of the cash transfers. The results also reveal that a higher proportion of households in SNNP region spent cash transfers in almost all categories. Particularly, the proportions that spent on non-food items or that saved are considerably higher. Consistent with this, the share of cash transfers used to buy non-food items, saved, and used to help other households is higher in SNNP while the share of cash transfers spent on

food is low. Statistical tests indicate that all of the differences indicated are statistically significantly different from zero.

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

| | Households that used cash transfers (%) | | | Cash transfers used (%) | | |
|---|---|---|---|---|---|---|
| | Amhara | Oromia | SNNP | Amhara | Oromia | SNNP |
| Food | 92.3 | 92.1 | 94.5 | 76.0 | 77.5 | 69.1 |
| Non-food | 38.9 | 39.9 | 52.5 | 15.7 | 14.6 | 19.0 |
| Crop production | 5.3 | 6.3 | 8.2 | 2.4 | 2.1 | 2.6 |
| Saving | 7.7 | 8.2 | 15.8 | 2.6 | 1.6 | 3.6 |
| Other HHs | 8.7 | 7.5 | 9.9 | 3.0 | 2.6 | 4.3 |
| Others | 1.9 | 7.0 | 6.6 | 0.2 | 1.6 | 1.3 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

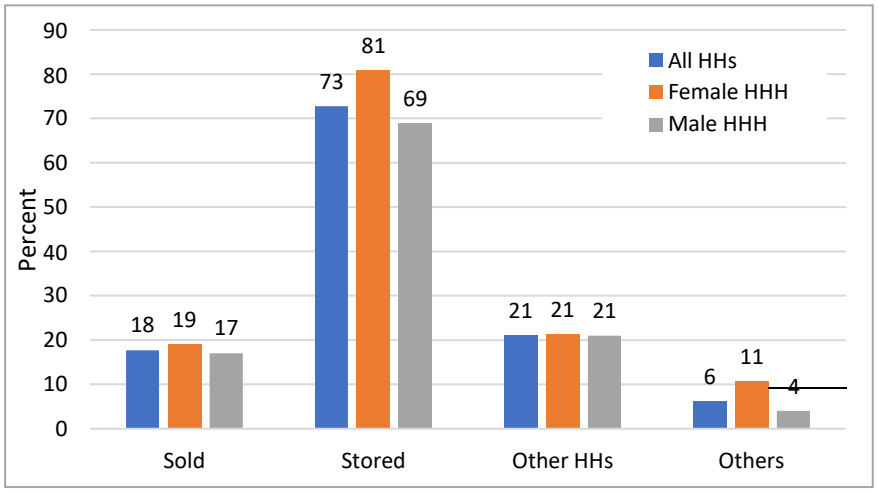
### 3.3 Food transfers

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 the percentages of households 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 of food transfers adds up to a 100 percent. These are summarized in Figures 3.8-3.10 and the second panel in Table 3.2.

Out of households that received food transfers nearly 18 percent sold food while 73 percent have food in storage at the time of the survey (Figure 3.6). Six percent of the households use food transfers for other purposes. The data indicate that no household gave food transfers to anyone in a position of authority. Given that the 'Others' category is composed of transfers given to people in a position of authority and to persons designated to collect the transfers, this implies six percent of the households gave food transfers for persons designated to collect the transfers.[8] No statistically significant differences exist in the proportion of female and male headed households that used food transfers for different purposes.

[8] Different classes of individuals were designated to collect transfers by such households - 31.6 percent were male household members, 42 percent female household members, 15.8 percent relatives who are not household members, 5.3 percent were neighbours, and 5.3 percent others.

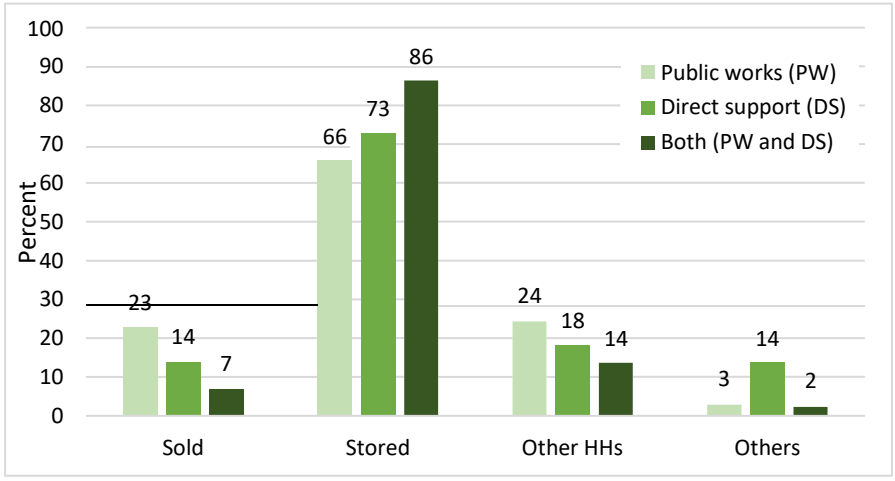
**Figure 3.6: Households that used food transfers for different purposes (%), by gendered household type**




***Note***: HHs = households; HHH = Headed household.
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

A relatively higher proportion of households that participated in both public works and direct support have food in their storage and the proportion of public works participant households that stored food is the lowest. The latter difference is statistically significant. The reverse holds when considering the proportion of households that sold food. That is, a statistically significantly higher proportion of households that participated in both public works and direct support sold food. The proportion of households that shared food transfers with other households and those that used food transfers for other purposes (gave food transfers for persons designated to collect the transfers) are not statistically significantly different across PSNP status of households.

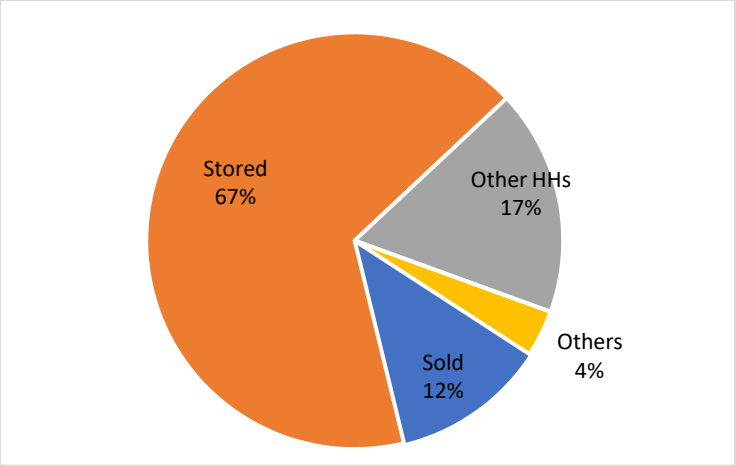
**Figure 3.7: Households that used food transfers for different purposes (%), by PSNP status**




***Note***: HHs = households
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Most of the grain transfers were in storage at the time of the survey (Figure 3.8), which is consistent with the depiction in Figure 3.6. Comparing the share of food transfers given to other households (17 percent) with the share of cash transfers given to other households (3 percent) indicates that households are more likely to provide in-kind assistance than cash transfers for other households.

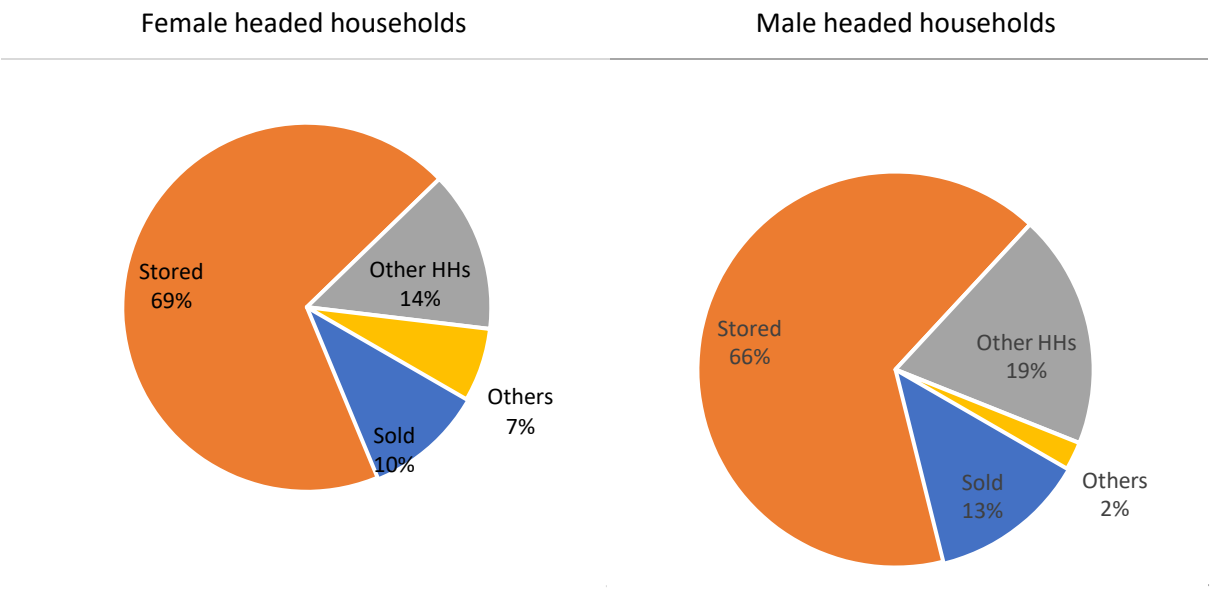
**Figure 3.8: Percentage of food transfers used for different purposes,**




***Note***: HHs = households
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Relative to male headed households, female headed households give a higher share of their food transfers for persons designated to collect their transfers and a lower share to help other households. Moreover, female headed households have a slightly higher food transfers in their storage and sold a slightly lower share. However, none of these differences are statistically significantly different from zero.

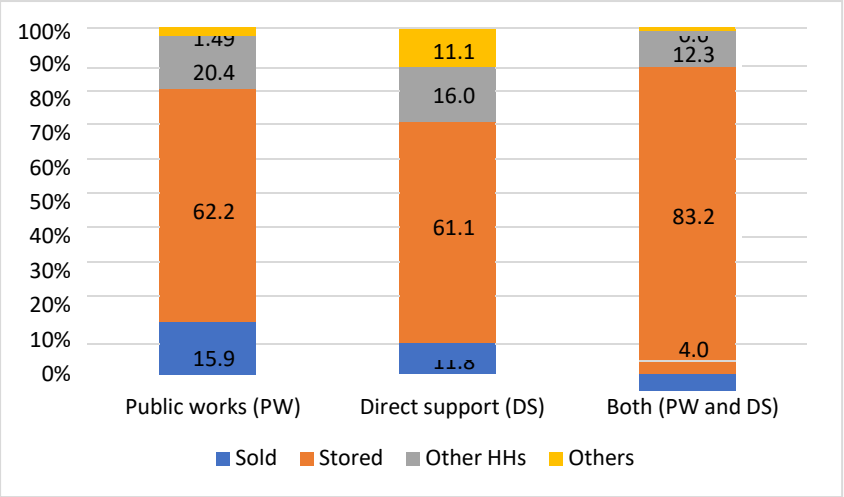
**Figure 3.9: Percentage of food transfers used for different purposes, by gendered household type**




***Note***: HHs = households
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

Households participating in direct support gave over 11 percent of their food transfers for persons designated to collect their payments and this (11 percent) is more than seven times the second highest share (1.5 percent) given by households participating in public works. However, the latter difference is not statistically significantly different from zero. Households participating in both public works and direct support have 83 percent of their food transfers in storage, which is over 20 percent higher than the share in storage for households participating in public works (61 percent), and this difference is statistically significant. In contrast, households participating in both public works and direct support sold a lower share of their food transfers, and this difference is statistically significant.

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




***Note***: HHs = households
***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

In Table 3.2 we summarize the proportion of households that used food transfers for different purposes and the importance of these purposes in total food transfers across regions. The summary indicates that the proportion of households that stored food transfers and the share of stored food are considerably higher in Amhara region than in Oromia and SNNP. In contrast, the proportion that sold food and the share of food sold is relatively higher in Oromia. A relatively higher proportion of households in SNNP used food transfers for other purposes. All of the differences indicated are statistically significantly different from zero excluding the proportion of households that used and the share of food transfers used for other purposes.

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

| | Households that used food transfers (%) | | | Food transfers used (%) | | |
|---|---|---|---|---|---|---|
| | Amhara | Oromia | SNNP | Amhara | Oromia | SNNP |
| Sold | 9.1 | 26.0 | 0.0 | 6.0 | 17.8 | 0.0 |
| Stored | 87.9 | 61.0 | 50.0 | 81.5 | 55.1 | 50.0 |
| Other HHs | 12.1 | 28.6 | 25.0 | 9.9 | 23.7 | 25.0 |
| Others | 4.5 | 6.5 | 25.0 | 2.5 | 3.4 | 25.0 |

***Source***: Authors' computation using data from the PSNP4 End-line Household Survey 2021.

### 3.4 Summary

The results indicate that over 90 percent of the households in highland areas spend cash transfers to buy food and 44 percent to buy non-food items. The results also indicate that households use nearly three-quarters of cash transfers to buy food and about 17 percent to buy non-food items. The share of cash transfers used for each of the remaining purposes is less than four percent. About 67 percent of the food transfer is in storage and 12 percent has been sold. Nearly 18 percent was given to other households, which compared to the 3 percent cash transfers given to other households, indicates that households are more likely to provide in-kind-assistance than cash-assistance for other households.

# Chapter 4: Trends in Food Security and Dietary Diversity

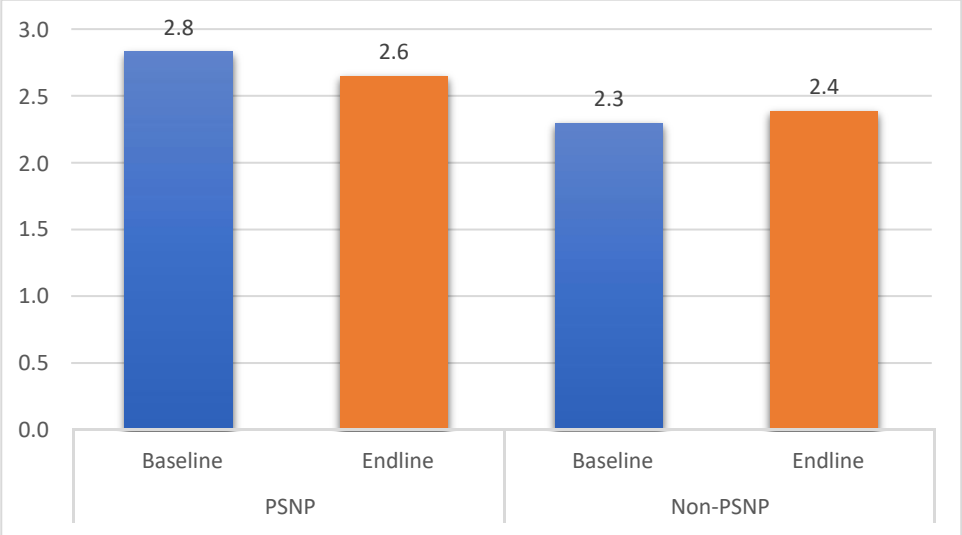
## 4.1 Introduction

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, consumption expenditures, and household level dietary diversity, comparing mean outcomes between baseline and end-line. Second, it summarizes the impact of the PSNP on these outcomes. As in previous rounds, food insecurity is measured by food gap, a self-reported measure at the household level that relies on counts of the 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, wereport the descriptive results showing trends between baseline and end-line. We also disaggregate someof 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.

## 4.2 Trends in food security

We begin by presenting the average food gap by PSNP status and survey round. As discussed above, food gap is calculated by taking the average of the response to the question: "How many months in the last 12 months did you have problems satisfying the food needs of the household?". The same question was administered at baseline and end-line for both PSNP and non-PSNP households. Note that a reduction in average food gap means an improvement in food security. Figure 4.1 reports the mean food gap by PSNP status between baseline and end-line. Table 4.5 and 4.6 also provide the statistical significance of the mean differences between PSNP and non-PSNP households.

**Figure 4.1: Mean food gap by PSNP status and round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

First, compared to non-PSNP households, PSNP households have experienced higher levels of mean food gap both at baseline and end-line. Second, between baseline and end-line, mean food gap has decreased by 0.2 months for PSNP households but increased by 0.1 months for non-PSNP households (Figure 4.1). Note that while the mean difference in food gap between PSNP and non-PSNP households at end-line is statistically significant at the 5% significance level, the mean difference in food gap among PSNP and non-PSNP households between baseline and end-line is only weakly statistically significant at 10 percent significance level (See Appendix Table 4.5 and Table 4.6).

We further disaggregate these results by the gender of the household head and PSNP status. Figure 4.2 (a) and (b) provide mean food gap for male and female headed households for PSNP and non-PSNP households.

**Figure 4.2(a): Mean food gap among PSNP households by gender of head and round (number of months)**

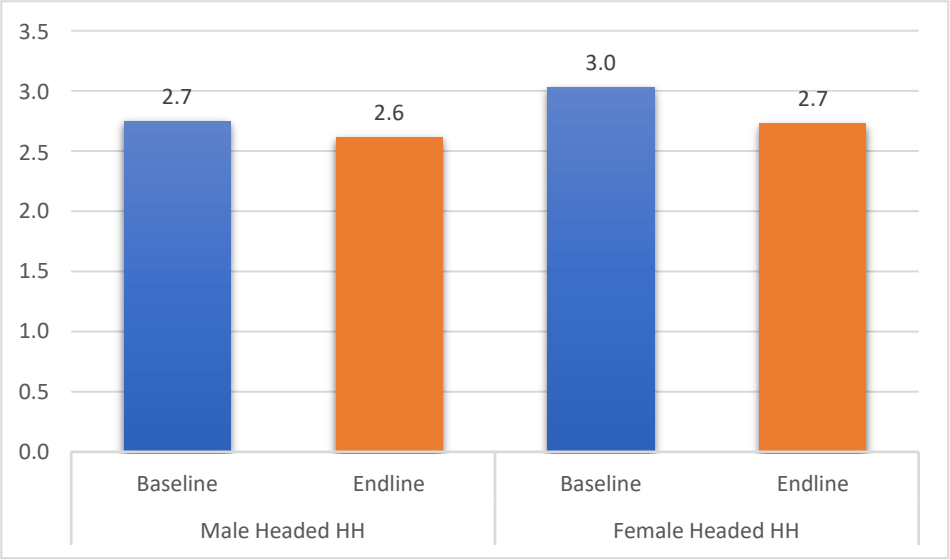
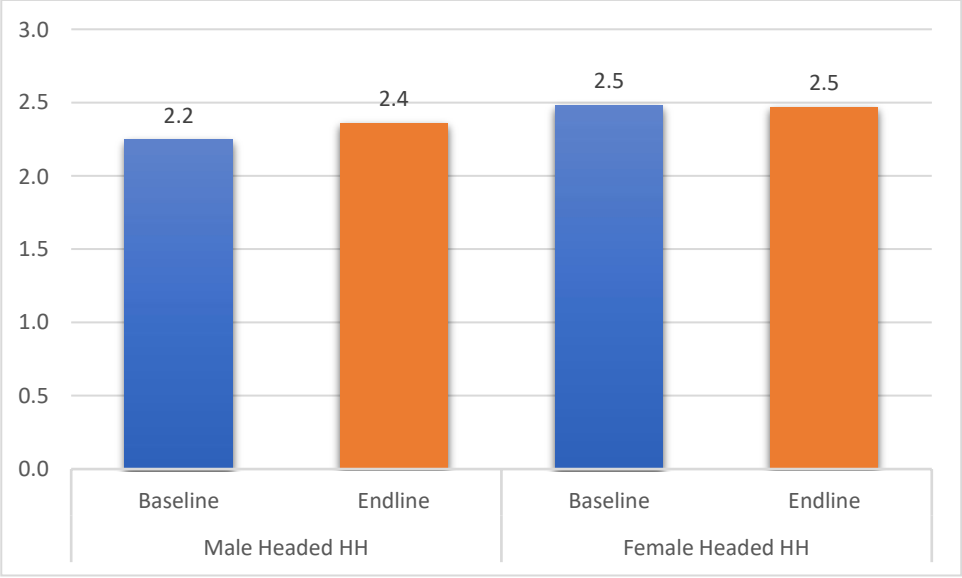



**Figure 4.2(b): Mean food gap among non-PSNP households by gender of head and round**




***Note***: HHs = household
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

There are two important highlights to note from these figures. First, female headed households tend to report higher mean food gap, regardless of their PSNP status. Second, mean food gap among PSNP female headed households has declined by 0.3 months (i.e., 9 days) between baseline and end-line (the difference is statistically significant at 10 percent significance level) but remained unchanged among the non-PSNP female headed households. On the other hand, mean food gap among non-PSNP households has increased by 0.2 months (6 days). This difference is not statistically different from zero even at 10 percent significance level.

**Figure 4.3(a): Mean food gap among PSNP households by region and round (number of months)**

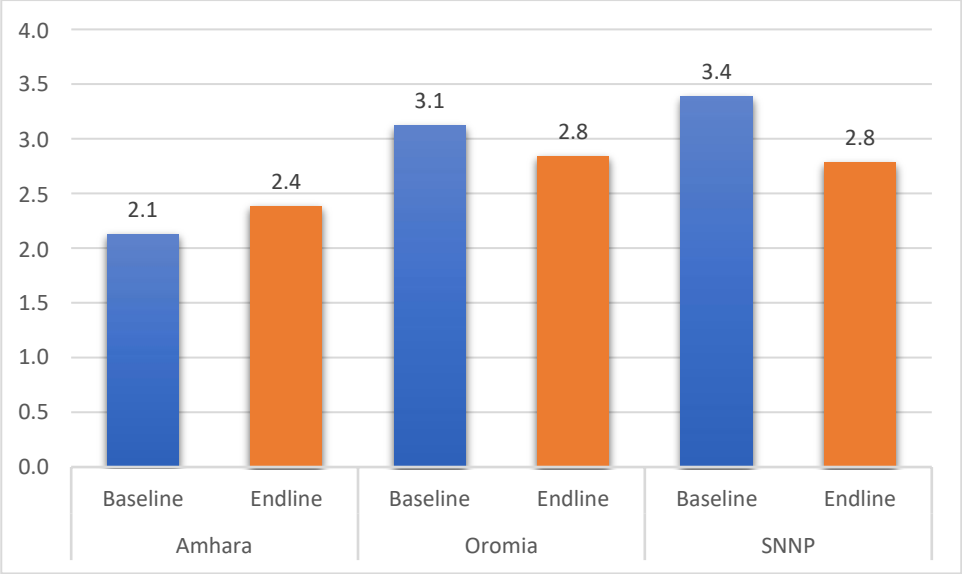
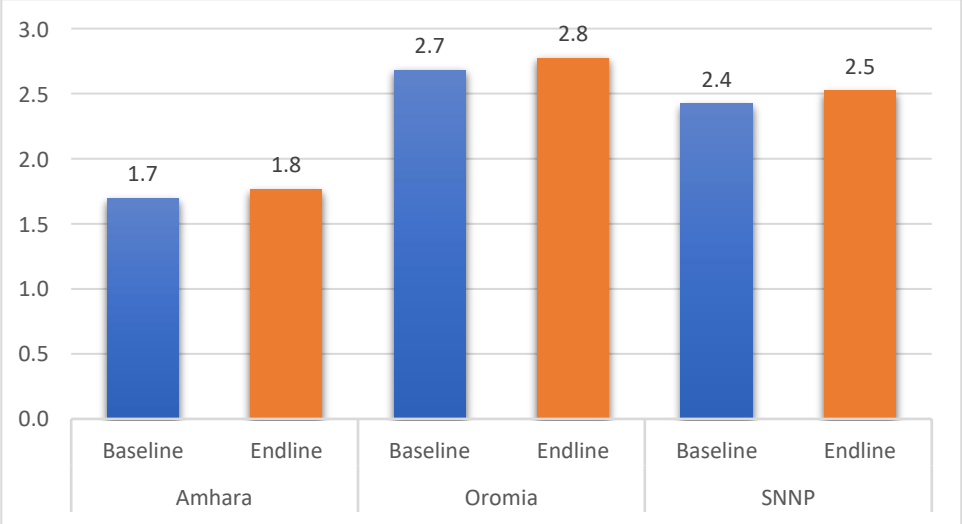



**Figure 4.3(b): Mean food gap among non-PSNP households by region and round (number of months)**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

To shed light on how households from each of the three regions fare in terms of food gap, in Figure 4.3 (a) and (b), we further disaggregate mean food gap by PSNP status, region, and survey round. Overall, between baseline and end-line, mean food gap among non-PSNP households has slightly increased in all

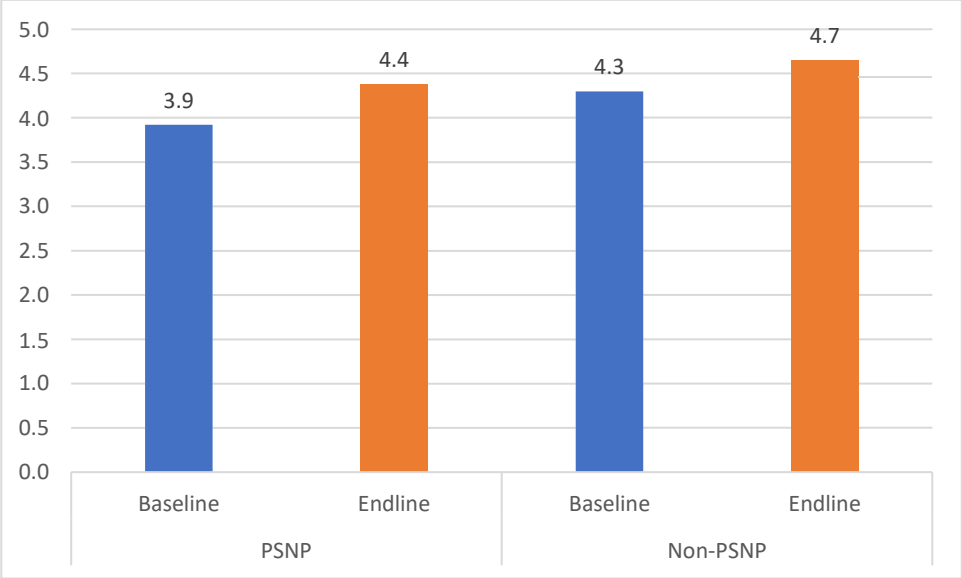
regions. In contrast, for the PSNP households mean food gap has declined for Oromia (by 9 days) and SNNP (by 18 days) but increased among PSNP households in Amhara by 9 days.

### 4.3 Trends in household dietary diversity

While the food gap measure indicates the extent to which the household has been able to fulfill its food needs, it does not inform about the quality or diversity of foods consumed. The diet diversity measure brings this *quality* dimension of food security. To assess this outcome, a diet diversity score is constructed using the household consumption module that captures which food groups are consumed by the household in a given time frame. 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. A value of one is assigned if the household has ever consumed in a given time frame from the 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.

Figure 4.4 reports mean household dietary diversity score (HDDS) by PSNP status and survey round. Table 4.5 in the appendix also provides t-tests of mean differences between PSNP and non-PSNP households at end-line; and Table 4.6 gives t-tests between baseline and end-line dietary diversity mean differences for PSNP and non-PSNP households. Consistent with the baseline findings, HDDS is generally low

**Figure 4.4: Mean household dietary diversity score by PSNP status and round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

Table 4.5 also show, at end-line, PSNP households have reported a marginally higher improvement (a

normalized mean difference of 0.1 more food groups, on average) in HDDS than the non-PSNP households and these differences are statistically highly significant at the 99 percent confidence level (Table 4.5 and 4.5)[9].

**Figure 4.5(a): Mean dietary diversity score among PSNP households by gender of head and round**

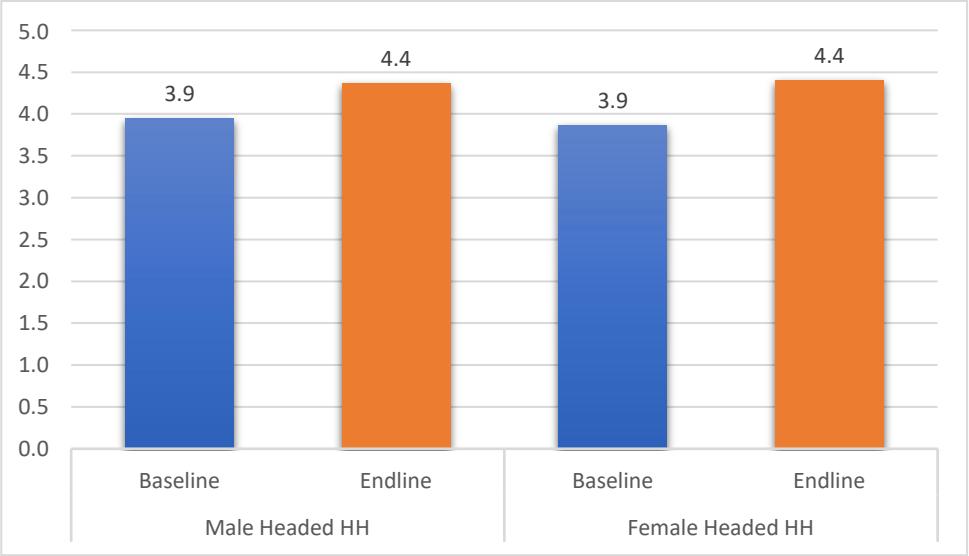
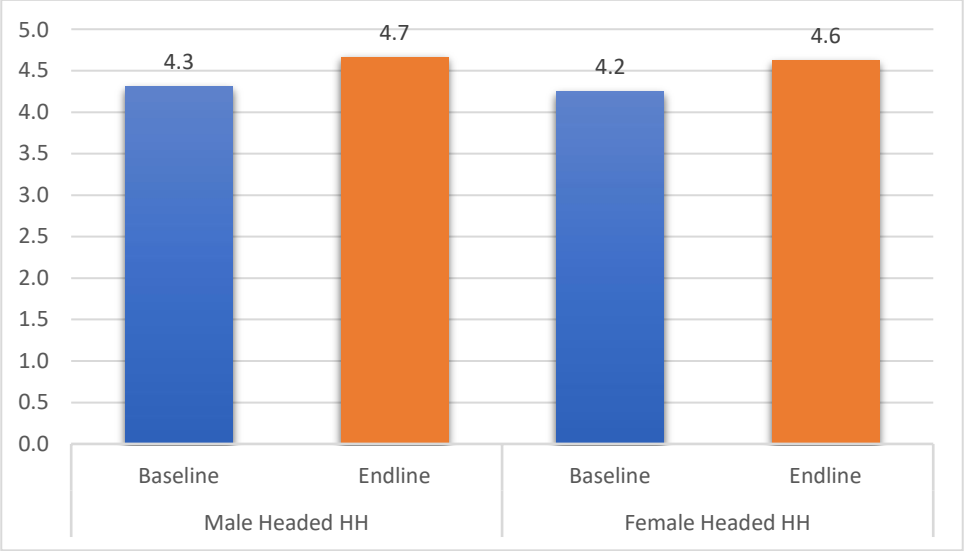



**Figure 4.5 (b): Mean dietary diversity score among non-PSNP households by gender of head and round**




***Note***: HHs = household
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

Figure 4.5 (a) and (b) disaggregate the mean household dietary diversity score by male and female

[9]Lower mean differences than the ones here are reported in Table 4.5 and Table 4.6 as they are normalized differences to account for sample variances (see notes under tables). Normalization takes dividing the standard deviation by its mean to make the standard deviation independent of units.

household heads, PSNP and non-PSNP, and survey rounds.

Three key findings from this gender disaggregation are: i) the overriding picture of the low mean HDDS remains regardless of the gender of the head of household. However, we also see improvements between baseline and end-line for all household groups regardless of PSNP status and gender of the head. Note that the mean difference between male headed and female headed households, as well as the changes over time are statistically significant at the 5 percent significance level; ii) female headed households score consistently lower mean HDDS than their male headed counterparts in the non-PSNP sample although the difference is statistically not different from zero at the 10 percent significance level; iii) on average, female headed households in the PSNP sample have relatively lower mean HDDS than female headed households in the non-PSNP sample. However, the improvement in mean HDDS is slightly higher for the PSNP than for the non-PSNP female headed households. All these latter mean differences are not statistically significant at the 10 percent significance level.

**Figure 4.6(a): Mean household dietary diversity score among PSNP households by region and round**

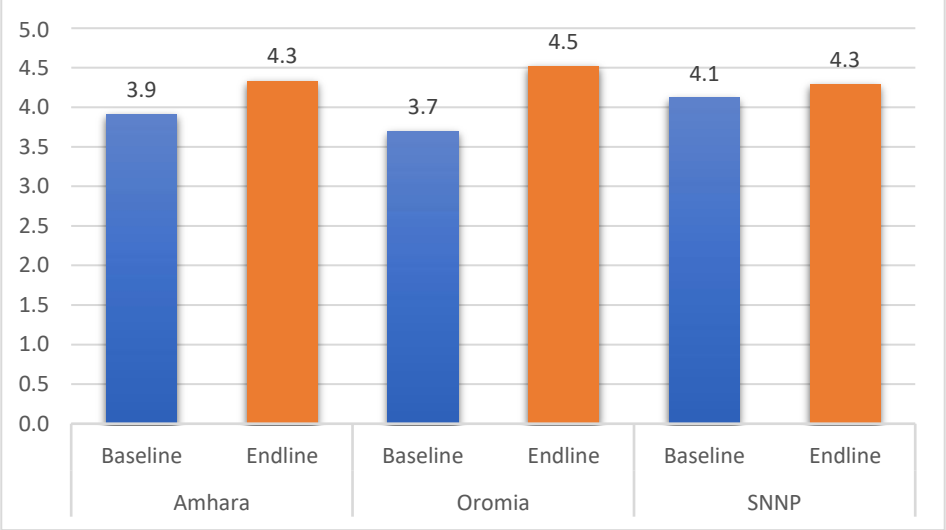
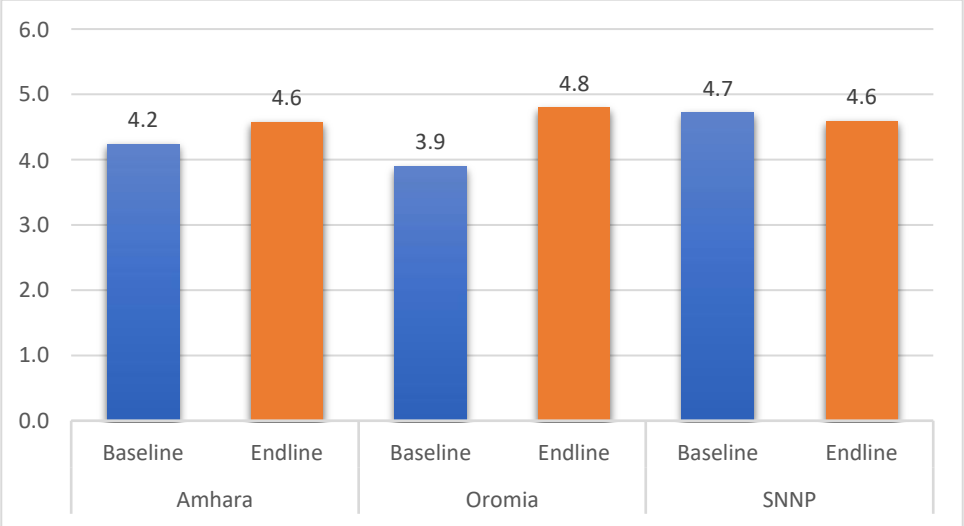



**Figure 4.6(b): Mean household dietary diversity score among non-PSNP households by region and round**



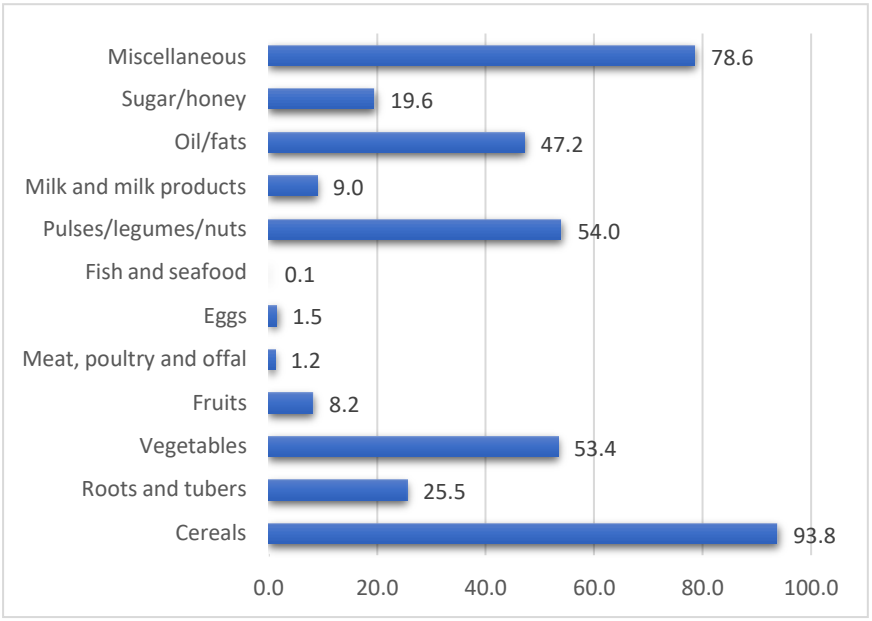

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

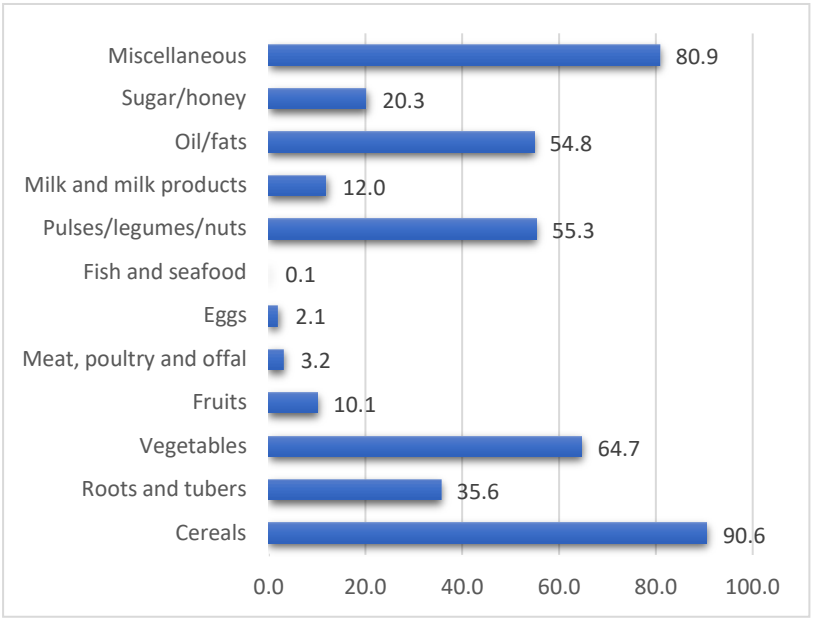
Regional disaggregation of these data, given in Figure 4.6(a) and 4.7 (b) also depict similar trends. Mean HDDS have increased in all regions, regardless of PSNP status, except for non-PSNP households in SNNP, where it has declined slightly (by 0.1 food groups, on average) (Figure 4.6(b)). All mean differences are statistically significant at the 5 percent significance level except those for SNNP.

Overall, we noted that HDDS is low despite slight improvements at end-line. This leaves us with the natural question, among the 12 food groups listed earlier, which of the food groups have households consumed from in the last 7 days prior to the survey day? Figure 4.7 summarizes proportion of (a) PSNP and (b) non-PSNP households consuming from each of the food groups mentioned above. Again, consistent with findings at baseline, the differences between PSNP and non-PSNP households are negligible. As can be seen from these figures, more than 90 percent of households have consumed cereals. The majority of households (more than 50 percent) have consumed vegetables, pulses and oils. On the contrary, consumption of animal source foods (meat, eggs, fish and seafoods, and dairy) and fruits were consumed by small proportions of PSNP and non-PSNP households.

**Figure 4.7: Percent of households consuming from different food groups at end-line by PSNP status[10]**

**a) PSNP households** **b) non-PSNP households**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

[10] Miscellaneous food group include spices, herbs, coffee, tea, diet soft drinks, and so on. Also note that we did not include the baseline to save space (also reported at midline) because these patterns remain monotonously similar across rounds and households' groups.

### 4.4 Trends in consumption expenditures

This section presents trends in household consumption expenditures in the Highlands. Total expenditure is sum of food and non-food expenditure measured for the last month preceding the survey. Values of 2021 are expressed in 2016 prices to adjust for inflation. We calculate per capita monthly food, non-food, and total expenditures by dividing these values to family size. Table 4.1 presents the mean per capita monthly consumption on food, non-food, and total expenditures by PSNP status and survey year. After adjusting for inflation, over the five years, we see that mean (real) total per capita monthly expenditures for all households has increased by 33 percent (last row of Table 4.1). This increase comes from the increase in per capita month food expenditure, which rose by 39 percent over the same period. Per capita monthly non-food expenditure has rather declined by 4 percent. There is negligible difference between PSNP and non-PSNP households in terms of the mean growth rate in real per capita consumption between 2016 and 2021. In terms of levels, PSNP households have slightly lower real per capita monthly expenditures than non-PSNP in both rounds. Differences between PSNP and non-PSNP in real food, non-food and total expenditure levels for 2021 is given in Table 4.7 (appendix) – and mean differences are statistically significant at 5 percent level.

**Table 4.1: Real per capita consumption expenditures (Birr in 2016 prices), by PSNP status and survey year**

| | 2016 | 2021 | Change (%) |
|---|---|---|---|
| **Non-PSNP households** | | | |
| Per capita monthly food expenditure | 275.7 | 382.6 | 38.8% |
| Per capita monthly non-food expenditure | 79.1 | 75.9 | -4.0% |
| Per capita monthly total expenditure | 345.9 | 458.5 | 32.6% |
| **PSNP households** | | | |
| Per capita monthly food expenditure | 247.0 | 341.2 | 38.1% |
| Per capita monthly non-food expenditure | 60.4 | 58.2 | -3.8% |
| Per capita monthly total expenditure | 298.3 | 399.4 | 33.9% |
| **All households** | | | |
| Per capita monthly food expenditure | 261.6 | 362.2 | 38.5% |
| Per capita monthly non-food expenditure | 69.9 | 67.2 | -3.9% |
| Per capita monthly total expenditure | 322.5 | 429.4 | 33.2% |

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.
***Note***: Values of 2021 are expressed in 2016 prices to adjust for inflation

Table 4.2 presents regional disaggregation of the mean pre capital real monthly food, non-food and total expenditures by PSNP status. Table 4.3 presents the same items in median terms and generally depicts similar patters as the mean. Clearly, there are important differences between regions in these consumption expenditures levels and growth rates. Four important regional differences can be drawn from Table 4.2. First, overall, households from Amhara have higher mean level of per capita monthly food expenditures. Second, except in Oromia, both at baseline and end-line, PSNP households have

consistently lower level of mean per capita monthly food expenditures than non-PSNP households. Third, growth rates of mean per capita monthly food expenditures for all households is higher in Amhara, followed by SNNP. Oromia has the lowest growth rate in mean per capita food expenditure. Fourth, Amhara has also positive growth rate of mean non-food per capita expenditure and Oromia and SNNP have seen declines in mean per capita non-food expenditures.

**Table 4.2: Mean real per capita consumption expenditures (Birr in 2016 prices), by region and survey year**

| | Amhara | | | Oromia | | | SNNP | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2016 | 2021 | Change (%) | 2016 | 2021 | Change (%) | 2016 | 2021 | Change (%) |
| **Non-PSNP households** | | | | | | | | | |
| Per capita monthly food expenditure | 310.9 | 485.6 | 56.2% | 262.6 | 312.6 | 19.1% | 252.9 | 338.4 | 33.8% |
| Per capita monthly non-food expenditure | 88.7 | 97.7 | 10.1% | 68.3 | 59.7 | -12.5% | 77.7 | 67.5 | -13.1% |
| Per capita monthly total expenditure | 397.3 | 583.3 | 46.8% | 307.9 | 372.4 | 20.9% | 325.9 | 405.9 | 24.6% |
| **PSNP households** | | | | | | | | | |
| Per capita monthly food expenditure | 260.3 | 407.5 | 56.5% | 265.6 | 279.1 | 5.1% | 217.0 | 317.8 | 46.5% |
| Per capita monthly non-food expenditure | 63.2 | 74.6 | 18.1% | 67.4 | 47.6 | -29.4% | 51.4 | 48.3 | -6.1% |
| Per capita monthly total expenditure | 319.2 | 482.1 | 51.0% | 311.4 | 326.7 | 4.9% | 263.5 | 366.1 | 39.0% |

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.
***Note***: Values of 2021 are expressed in 2016 prices to adjust for inflation

**Table 4.3: Median real per capita consumption expenditures (Birr in 2016 prices), by region and survey**

| | Amhara | | | Oromia | | | SNNP | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2016 | 2021 | Change (%) | 2016 | 2021 | Change (%) | 2016 | 2021 | Change (%) |
| **Non-PSNP households** | | | | | | | | | |
| Per capita monthly food expenditure | 227.1 | 301.2 | 32.6% | 192.2 | 211.6 | 10.1% | 176.2 | 187.1 | 6.2% |
| Per capita monthly non-food expenditure | 57.0 | 58.3 | 2.3% | 48.3 | 34.5 | -28.7% | 47.6 | 37.7 | -20.8% |
| Per capita monthly total expenditure | 300.3 | 386.2 | 28.6% | 237.9 | 263.8 | 10.9% | 234.9 | 253.5 | 7.9% |
| **PSNP households** | | | | | | | | | |
| Per capita monthly food expenditure | 195.3 | 270.3 | 38.4% | 195.6 | 175.4 | -10.3% | 154.7 | 181.4 | 17.3% |
| Per capita monthly non-food expenditure | 38.6 | 44.5 | 15.3% | 46.7 | 28.8 | -38.3% | 31.0 | 29.6 | -4.6% |
| Per capita monthly total expenditure | 248.5 | 342.7 | 37.9% | 226.8 | 208.5 | -8.1% | 199.4 | 226.1 | 13.4% |

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.
***Note***: Values of 2021 are expressed in 2016 prices to adjust for inflation

## 4.5 Impact of the PSNP on food security, household dietary diversity, and consumption expenditures

Table 4.4 presents the impact of the PSNP4 on food security, dietary diversity and changes in real per capita consumption expenditures based on the matching techniques. For the full sample, we find that the PSNP reduces food gap, on average, by 0.18 month (or 6 days per year) and the estimated impact is statistically significant at 10 percent significance level. However, the PSNP did not have statistically significant impact on household dietary diversity score, and three consumption expenditure measures, namely change in per capita food, change in per capita non-food, and change in per capita total expenditures. In other words, we cannot reject the hypothesis that the impact of the PSNP on these latter outcomes is not statistically different from zero. 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) in 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 4.4: Impact of PSNP4 on food security and consumption expenditures, full sample.**

| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| Change in food gap | -0.177* | 4,921 | 0.089 |
| | (0.100) | | |
| Change in Household Dietary Diversity Score (HDDS) | 0.059 | 4,924 | 0.354 |
| | (0.073) | | |
| Change in per capita real total consumption expenditures | -30.186 | 4,509 | 112.503 |
| | (19.054) | | |
| Change in per capita real total food expenditures | -24.768 | 4,353 | 109.398 |
| | (18.601) | | |
| Change in per capita real total nonfood expenditures | -4.137 | 4,509 | -3.218 |
| | (3.722) | | |

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.
***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$; Results based on Nearest Neighbor Matching

## 4.6 Summary

This chapter presented trends in household food security, consumption expenditures, and dietary diversity and reported the impact estimates of the PSNP on these outcomes.

- The average PSNP household reports higher levels of food gap than non-PSNP households both at baseline and end-line. However, between baseline and end-line, mean food gap has decreased by 0.2 months for PSNP households but increased by 0.1 months for non-PSNP households.

- While female headed households report higher mean food gap, regardless of their PSNP status, mean food gap among female headed PSNP households has declined by 0.3 months between baseline and end-line but remained unchanged among the non-PSNP female headed households.
- Regional disaggregation suggests mean food gap has slightly increased for non-PSNP households in all regions. In contrast, it has declined for PSNP households in Oromia and SNNP but slightly increased in Amhara by about 9 days per year.
- Consistent with previous assessments on the same sample, HDDS is dismally low among households in this context and, on average, households in these areas consume only from 4.5 food groups out of the possible 12 food groups at end-line. This has shown some improvements from baseline by an average of 0.4 food group for all households.
- Female headed households in the PSNP sample have relatively lower mean HDDS than female headed households in the non-PSNP sample. However, the improvement in mean HDDS is slightly higher for the PSNP than for the non-PSNP female headed households.
- Overall, between baseline and end-line, mean per capita monthly total expenditures for all households has increased by 33 percent (last row of Table 4.1). This increase comes from the increase in per capita monthly food expenditures, which rose by 39 percent over the same period. Per capita monthly non-food expenditure has rather declined by 4 percent.
- There is negligible difference between PSNP and non-PSNP households in terms of the mean growth rate in real per capita consumption expenditures between 2016 and 2021. In terms of levels, PSNP households have slightly lower real per capita monthly expenditures than non-PSNP in both rounds. We note that these latter mean differences between PSNP and non-PSNP are statistically significant at 5 percent level.
- There are important regional differences in mean per capita monthly food, non-food, and total consumption expenditure levels and growth rates. Overall, households in Amhara have higher level of mean per capita monthly food expenditures. Growth rates in mean per capita monthly food expenditures for all households is also higher in Amhara, followed by SNNP. Oromia has the lowest growth rate in mean per capita food expenditure. While Amhara has shown a positive growth rate in mean non-food per capita expenditures, Oromia and SNNP experienceddeclines in mean per capita non-food expenditures.
- Results from impact evaluation show that the PSNP has been able to reduce food gap, on average, by about 6 days per year but did not have statistically significant impact on household dietary diversity and on changes in real per capita monthly expenditures. We note possible explanations for these outcomes.

**Table 4.5: Food security outcome summary by the Treatment and Control samples in 2021**

| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Food gap | 2.647 | 2.340 | 2,235 | 2.384 | 2.566 | 2,691 | 0.000 | 0.107 |
| Household Dietary Diversity Score (HDDS) | 4.376 | 1.850 | 2,235 | 4.653 | 1.884 | 2,691 | 0.000 | -0.148 |

***Notes***: Data are from the 2021 round of the PSNP4. 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 and non-PSNP sample variances.

***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

**Table 4.6: Trend in Food security outcomes between the PSNP and non-PSNP samples, 2016-2021**

| | Baseline | | | End-line | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| **PSNP beneficiary household** | | | | | | | | |
| Food gap | 2.829 | 2.583 | 2,233 | 2.647 | 2.340 | 2,235 | 0.013 | 0.074 |
| Household Dietary Diversity Score (HDDS) | 3.921 | 1.823 | 2,235 | 4.376 | 1.850 | 2,235 | 0.000 | -0.248 |
| **Non-PSNP beneficiary households** | | | | | | | | |
| Food gap | 2.295 | 2.548 | 2,689 | 2.384 | 2.566 | 2,691 | 0.204 | -0.035 |
| Household Dietary Diversity Score (HDDS) | 4.297 | 1.876 | 2,691 | 4.653 | 1.884 | 2,691 | 0.000 | -0.190 |

***Notes***: Data are from baseline and end-line of the PSNP4 data. 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 baseline sample mean minus the end-line sample mean to the square root of one half of the sum of the baseline and end-line sample variances.

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

**Table 4.7: Consumption Expenditure outcome summary by the Treatment and Control samples in 2021**

| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Real total consumption expenditures | 1,688.005 | 1,670.041 | 2,223 | 2,041.051 | 2,035.797 | 2,291 | 0.000 | -0.190 |
| Real total food expenditures | 1,441.070 | 1,590.375 | 2,223 | 1,695.133 | 1,919.747 | 2,291 | 0.000 | -0.144 |
| Real total nonfood expenditures | 246.935 | 298.461 | 2,223 | 345.918 | 409.425 | 2,291 | 0.000 | -0.276 |
| Per capita real total consumption expenditures | 399.391 | 475.489 | 2,223 | 458.546 | 554.529 | 2,291 | 0.000 | -0.115 |
| Per capita real total food expenditures | 341.234 | 440.251 | 2,223 | 382.623 | 520.454 | 2,291 | 0.004 | -0.086 |
| Per capita real total nonfood expenditures | 58.157 | 105.653 | 2,223 | 75.923 | 113.057 | 2,291 | 0.000 | -0.162 |

***Notes***: Data are from the 2021 round of the PSNP4. 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

***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

# Chapter 5: Assets, Shocks, and Resilience

## 5.1 Introduction

This chapter presents descriptive results on trends between the baseline and end-line in asset holdings of PSNP and non-PSNP households. We focus on livestock assets and productive assets. Livestock asset is measured in Tropical Livestock Units (TLU)[11] as well as in values or monetary units. Monetary units are adjusted for inflation to make them comparable across periods. We specifically compare baseline (or 2016) mean outcomes against end-line (2020) 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. We begin with trends in livestock asset holdings measured in TLU and values.

## 5.2 Trends in livestock asset holdings

This section summarizes the mean changes between baseline and end-line in livestock asset holdings of PSNP and non-PSNP households. We further disaggregate these results by gender of the household head and region. Table 5.1 reports the trends in asset holdings mainly livestock by PSNP and survey round. We first focus on livestock assets using the TLU measure and values in Birr. It is shown that the average PSNP household owned 2 TLUs at baseline with a total average value of 6,200 Birr. This increases to an average of 2.5 TLUs valuing 14,716 Birr at end-line. These changes between baseline and end-line are statistically significant at the 1 percent significance level. A comparable non-PSNP household owned, on average, 3 TLUs with an average total livestock value of Birr 9,428 at baseline. These grew to 3.2 TLU with a household level average livestock value of Birr 19,400. The changes in these outcomes are again statistically significant at the 1 percent significance level.

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

**Table 5.1: Trend in asset holdings between the PSNP and non-PSNP samples, 2016-2021**

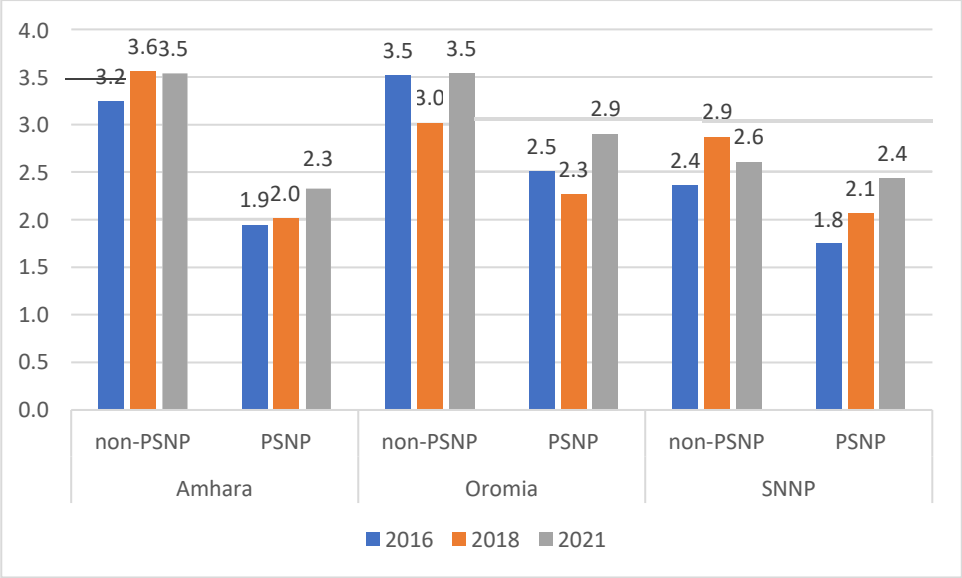
| | 2016 | | | 2021 | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **N** | **Mean** | **SD** | **N** | | |
| **PSNP beneficiary households** | | | | | | | | |
| Current total tropical livestock units | 2.0 | 2.4 | 2,234 | 2.5 | 3.1 | 2,235 | 0.000 | -0.178 |
| Current total value of livestock (Birr) | 6,195.5 | 7,849.1 | 2,234 | 14,726.3 | 18,682.0 | 2,235 | 0.000 | -0.595 |
| Owned by women: Current total value of livestock (Birr) | 555.1 | 2,431.5 | 1,383 | 2,253.1 | 35,413.4 | 1,529 | 0.075 | -0.068 |
| **Non-PSNP beneficiary households** | | | | | | | | |
| Current total tropical livestock units | 3.0 | 3.2 | 2,690 | 3.2 | 3.5 | 2,691 | 0.045 | -0.055 |
| Current total value of livestock (Birr) | 9,428.3 | 10,213.1 | 2,690 | 19,399.1 | 23,078.0 | 2,691 | 0.000 | -0.559 |
| Owned by women: Current total value of livestock (Birr) | 871.5 | 4,168.3 | 1,874 | 1,518.6 | 8,359.3 | 2,074 | 0.002 | -0.098 |

***Notes:*** Data are from baseline and end-line of the PSNP4 data. 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 baseline sample mean minus the end-line sample mean to the square root of one half of the sum of the baseline and end-line sample variances.

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

In Figure 5.1 we further disaggregate the TLU results by region, PSNP status and survey round[12]. There are two important observations from this graph. First, across all regions, compared to the non-PSNP, PSNP households own consistently lower average TLUs of livestock (the difference is statistically significant at the 5 percent significance level). However, growth rate in average TLU is higher among PSNP (an average growth of 0.5) than non-PSNP households (an average growth of 0.2 TLUs between 2016 and 2021). Second, in all regions, livestock ownership measured in TLU, has consistently increased across regions and regardless of PSNP status (except in the non-PSNP households in Oromia, which was already among the highest). SNNP has relatively noticeably lower average TLUs than the rest of the regions. Average TLUs in Oromia and Amhara are roughly comparable.

**Figure 5.1: Livestock holding in TLU by region, PSNP status, and round**

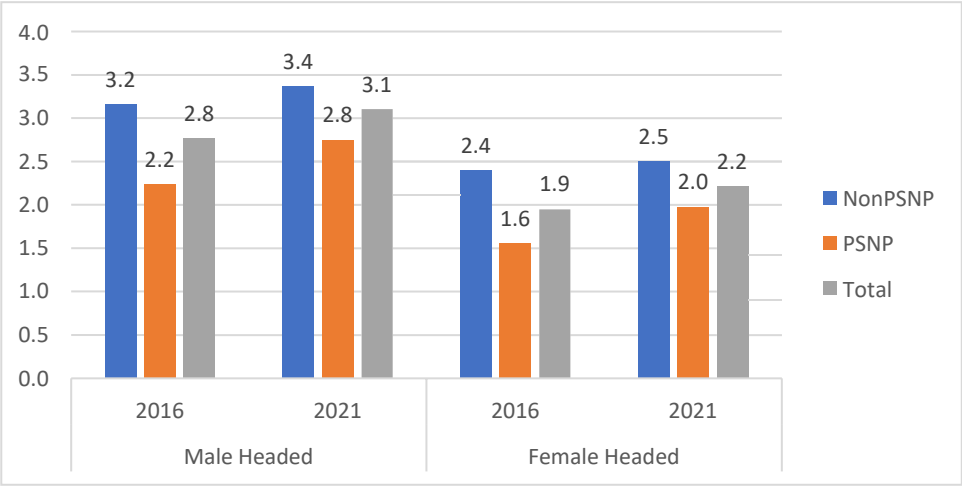



***Source***: Authors' computation using data from the PSNP IV Baseline (2016), Mid-line (2018), and End-line (2021) Household Surveys

Figure 5.2 breaks these households level ownership of livestock into male headed and female headed ownership. This disaggregation uncovers another important dimension of heterogeneity in livestock ownership between male and female headed households. Overall, female headed households are less endowed with livestock than their male counterparts and that female headed PSNP households are even less endowed with livestock as compared to non-PSNP female headed households. These differences are statistically significant at the 5 percent significance level.

[12] Midline averages are included for ease of reference. However, we note that there is slight difference in the definition of PSNP beneficiary status between Midline and End-line. At Midline, a household is a PSNP beneficiary based on PSNP status at Baseline. At End-line, a household is a PSNP beneficiary if she participated in the PSNP4 for either 3 or 4 years.

**Figure 5.2: Livestock holding in TLU by gender of head and PSNP status and round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

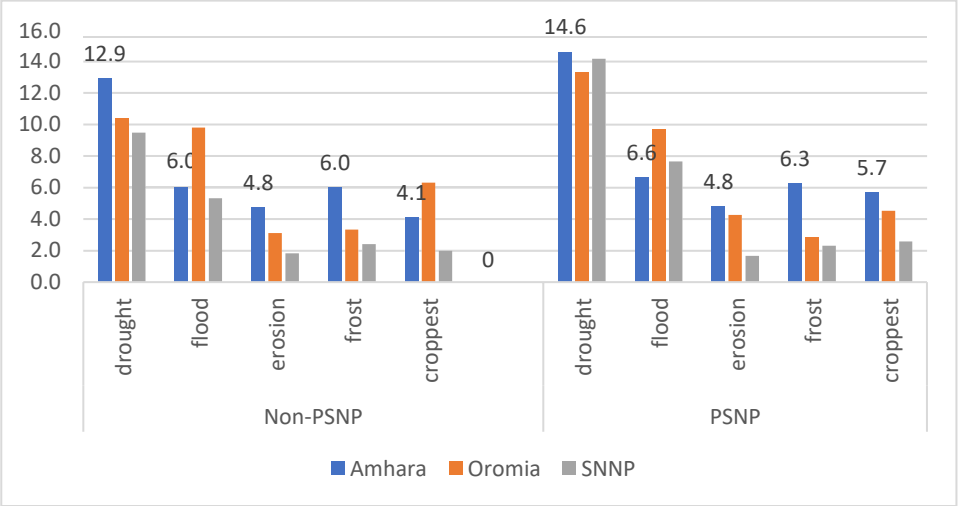
## 5.3 Shocks and resilience

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 subsequent sections (see impact estimates), 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.

### *5.3.1 Shocks*

To understand whether and what type of shocks households experience over a three-period horizon, we asked our respondent the following question: "Thinking about the last three years what were the most important shocks you experienced?" Figure 5.3 summarizes the proportion of households reporting each of these shock types by PSNP status and region. These are the five most important shocks experienced by most of the respondents. We note that there is difference between PSNP and non-PSNP households when it comes to the type of shocks faced. Further we note that drought is the most important shock for most households in all regions followed by floods and erosion. Drought shock is reported highest by households in Amhara, followed by Oromia and SNNP. Overall, about 15 percent of PSNP (13 percent of non-PSNP) households from Amhara report drought as first most important shock. Despite some variabilities, pests that attack crops and frost are among the least important shocks for all households in all regions.

**Figure 5.3: Percent of PSNP and non-PSNP households experienced shocks by region**




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

Do shocks vary by the gender of the household head? This is the next question we address using the same question described above. Table 5.4 (a) and (b) report the responses by gender of the household head and PSNP benefit. In almost all cases, PSNP households are more likely to report one of these shocks than non-PSNP households. And female headed households are slightly less likely to report one of these shocks. Other than these, both male and female headed households exhibit similar general patterns of shock experience, regardless of gender of head or PSNP status.

**Figure 5.4(a): Percent of households who experienced shocks by region and PSNP: Male head**

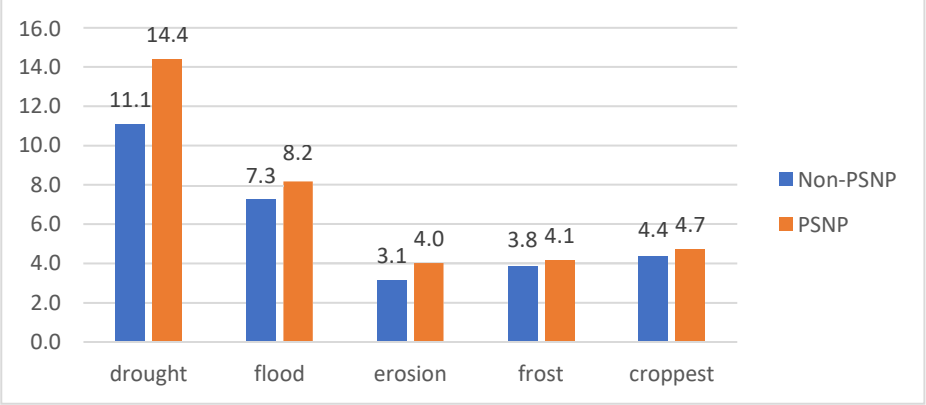
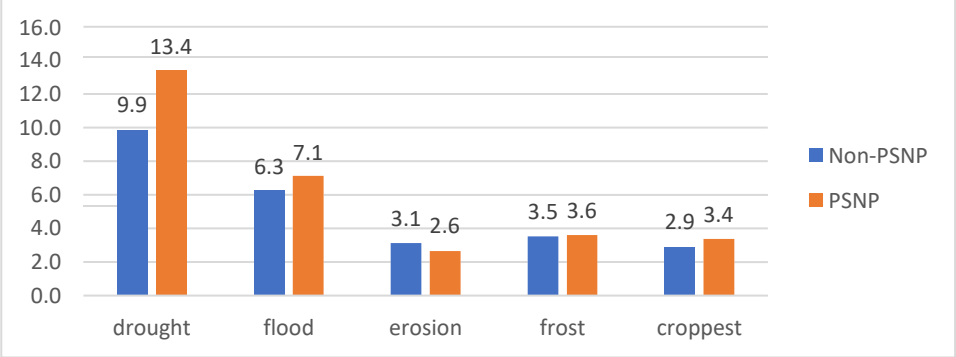



**Figure 5.4(b): Percent of households who experienced shocks by region and PSNP: Female head**




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

### *5.3.2 Distress asset sales*

How resilient are PSNP households to shocks and do they sell productive assets to cope with it? Does resilience vary by PSNP and non-PSNP or region where the household comes from? Table 5.2 shows 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 baseline and end-line, and region. This table shows the overall picture of changes in these outcomes over time by region. Overall, among the three assettypes, distress sales of livestock assets for food or emergency cash needs is reported by a relatively higher proportion of households (21 percent at end-line), with an increasing trend between baseline and end-line by 6 percentage points each for food and emergency cash needs (last column of Table 5.2). Among regions, Amhara and Oromia appear to be on the high side (24 percent each at end-line reporting sales of livestock for food needs). In terms of trends, the proportion of SNNP households reporting with distress sales of livestock for food has increased by 7 percentage points (the highest increase than any of the rest between baseline and end-line).

**Table 5.2: Percent of households reporting distress sales of assets in the last two years by round and region.**

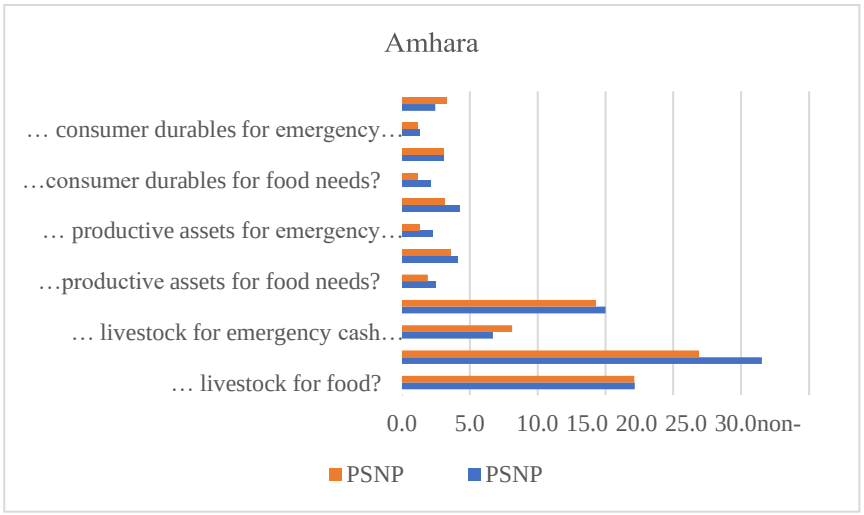
| Forced to sell any … | Region | Amhara | Oromia | SNNP | Total |
|---|---|---|---|---|---|
| … livestock for food needs? | Baseline | 17.1 | 20.0 | 8.4 | 15.0 |
| | End-line | 24.6 | 24.1 | 15.5 | 21.3 |
| … livestock for emergency cash needs? | Baseline | 7.4 | 11.7 | 7.4 | 8.7 |
| | End-line | 14.8 | 19.5 | 12.3 | 15.4 |
| …productive assets for food needs? | Baseline | 2.1 | 3.2 | 5.5 | 3.6 |
| | End-line | 3.8 | 3.1 | 7.6 | 4.9 |
| … productive assets for emergency cash needs? | Baseline | 1.7 | 3.5 | 4.4 | 3.2 |
| | End-line | 3.6 | 5.0 | 5.9 | 4.9 |
| …consumer durables for food needs? | Baseline | 1.6 | 3.5 | 4.3 | 3.1 |
| | End-line | 3.0 | 4.3 | 5.5 | 4.3 |
| … consumer durables for emergency cash needs? | Baseline | 1.2 | 2.6 | 2.9 | 2.2 |
| | End-line | 2.8 | 3.8 | 5.3 | 4.0 |

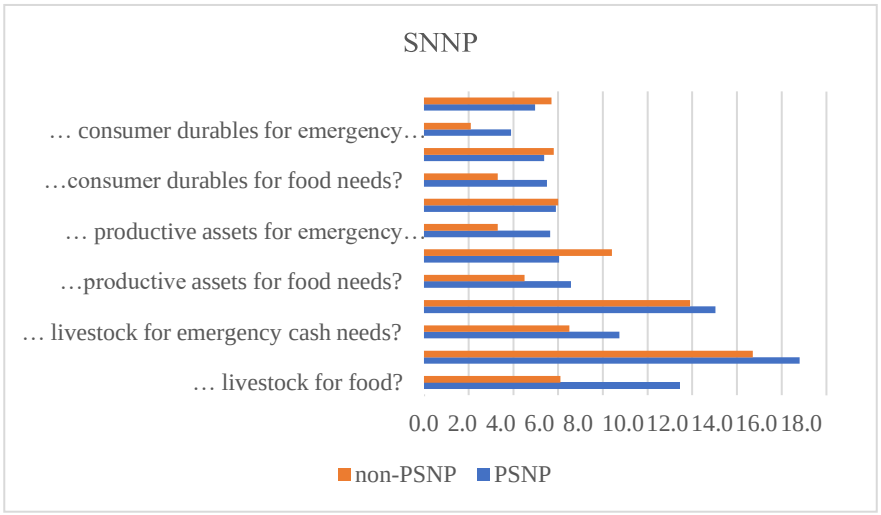
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

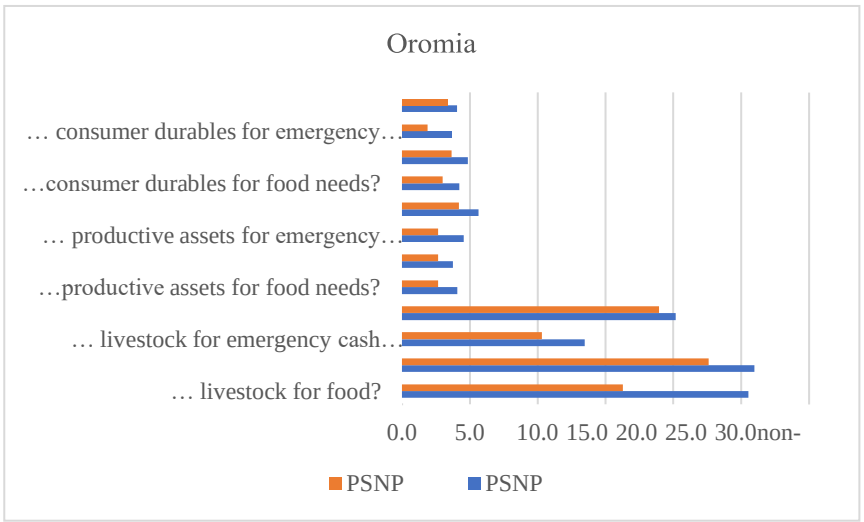
Figure 5.5 disaggregates these results by PSNP and non-PSNP for the baseline and end-line (the same results are also reported in Table 5.3, appendix). Clearly, in the two years preceding the end-line, distress sales of livestock assets for food and emergency cash needs were most common invariably among PSNP and non-PSNP households of all regions. However, distress sales of productive assets and consumer durables seem higher in SNNP than the rest two regions. Overall, PSNP households (23 percent at end-line) are more likely to report distress sales of livestock for food than non-PSNP households (20 percent at end-line). However, the increase in distress sales of livestock for food between baseline and end-line is higher for non-PSNP (7 percentage points) than for PSNP (5 percentage points). A similar trend is observed in distress sales of livestock for emergency cash needs. Among regions, PSNP households in Amhara (27 percent) and Oromia (26 percent) are more likely to report distress sales of livestock for food at end-line. Compared to livestock assets, distress sales of productive assets and consumer durables are less likely to be reported by both PSNP and non-PSNP households in all regions (Table 5.5). From the payments chapter, we know that PSNP households do not receive their complete entitlements and that payments are not timely or predictable. No

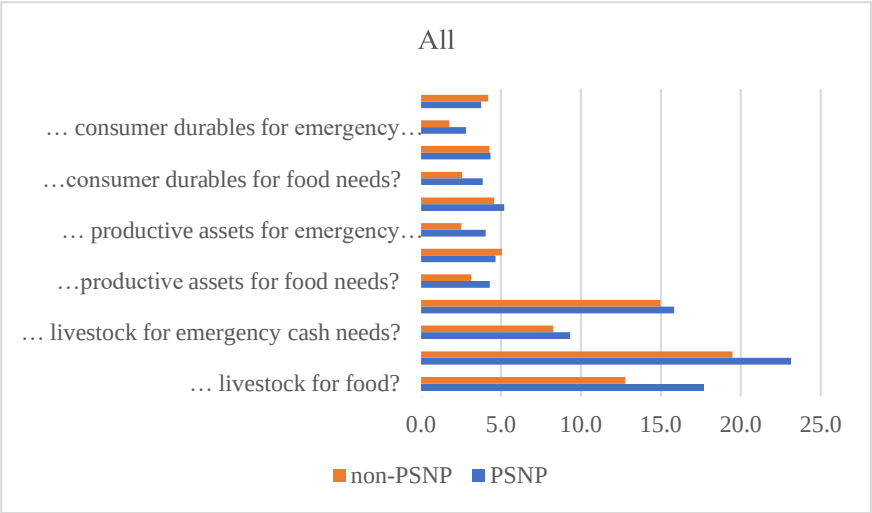
surprise, PSNP households are reporting distress asset sales when transfers are delayed or altogether incomplete.

**Figure 5.5: In the last two years (asked in April-May 2021), have you been forced to sell any of the following assets? (%)**






***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

**Table 5.3: Percent of households reporting distress sales of assets in the last two years by round, region, and PSNP status.**

| | | Amhara | | Oromia | | SNNP | | Total | |
|---|---|---|---|---|---|---|---|---|---|
| **Forced to sell any …** | | **PSNP** | **non-PSNP** | **PSNP** | **non-PSNP** | **PSNP** | **non-PSNP** | **PSNP** | **non-PSNP** |
| … livestock for food? | Baseline | 17.2 | 17.1 | 25.5 | 16.3 | 11.4 | 6.1 | 17.6 | 12.8 |
| | End-line | 26.5 | 21.9 | 26.0 | 22.6 | 16.8 | 14.7 | 23.1 | 19.5 |
| … livestock for emergency cash needs? | Baseline | 6.7 | 8.1 | 13.5 | 10.3 | 8.7 | 6.5 | 9.3 | 8.3 |
| | End-line | 15.0 | 14.3 | 20.2 | 19.0 | 13.0 | 11.9 | 15.8 | 15.0 |
| …productive assets for food needs? | Baseline | 2.5 | 1.9 | 4.1 | 2.7 | 6.6 | 4.5 | 4.3 | 3.1 |
| | End-line | 4.1 | 3.5 | 3.8 | 2.7 | 6.0 | 8.4 | 4.7 | 5.1 |
| … productive assets for emergency cash needs? | Baseline | 2.2 | 1.3 | 4.5 | 2.7 | 5.6 | 3.3 | 4.0 | 2.5 |
| | End-line | 4.2 | 3.2 | 5.6 | 4.2 | 5.9 | 6.0 | 5.2 | 4.6 |
| …consumer durables for food needs? | Baseline | 2.1 | 1.1 | 4.2 | 3.0 | 5.5 | 3.3 | 3.9 | 2.6 |
| | End-line | 3.1 | 3.0 | 4.9 | 3.7 | 5.4 | 5.8 | 4.3 | 4.3 |
| … consumer durables for emergency cash needs? | Baseline | 1.3 | 1.1 | 3.6 | 1.9 | 3.9 | 2.1 | 2.8 | 1.7 |
| | End-line | 2.5 | 3.3 | 4.1 | 3.3 | 5.0 | 5.7 | 3.8 | 4.2 |

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

### 5.4 Impact of the PSNP on household assets and resilience

The previous sections have presented the trends in assets, shocks and resilience outcomes. This section presents the impact of the PSNP on these outcomes. Table 5.4 reports the impact of the PSNP on livestock assets owned both by women and the household, as well as on distress sales of productive assets. Difference-in-differences with Nearest Neighbor Matching estimator is used in the estimation. Details of these are given in Chapter 2 (method chapter). Here we provide the key results.[13]

We find that the PSNP increased livestock assets measured by TLU. As can be seen from the first row of Table 5.4, the PSNP has increased TLU by 25 percent, on average. The estimate is statistically significant at the 5 percent significance level. However, in a rather awkward and contrary to the above finding, we find a negative and statistically significant estimate for total value of livestock assets. This result is counter intuitive and is perhaps due to measurement error. In this regard, the TLU measure, being less prone to measurement error, is a more reliable indicator of livestock assets.

On the other hand, the PSNP did not have any statistically significant impacts on any of the rest of the assets including changes in total value of livestock owned by women, changes in total value of productive assets, changes in income diversification, or on any of the resilience indicators (forced sales of livestock and productive assets for food and emergency cash needs).

**Table 5.4: The impact of the PSNP on assets and distress sales of assets, Nearest Neighbor Matching**

| | Treatment effect (SE) from NNM | N |
|---|---|---|
| Change in total tropical livestock units | 0.252** | 4,923 |
| | (0.105) | |
| Change in total value of livestock | -1,451.311** | 4,923 |
| | (572.555) | |
| Change in total value of livestock (women) | 950.033 | 3,060 |
| | (584.611) | |
| Forced to sell any productive assets for food needs | -0.005 | 4,924 |
| | (0.007) | |
| Forced to sell any productive assets for emergency cash needs | 0.007 | 4,924 |
| | (0.007) | |
| Change in total values of productive assets (Birr) | -153.664 | 4,924 |
| | (133.780) | |
| Change in income diversification, baseline to end-line | 0.022 | 4,924 |
| | (0.020) | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

The lack of impact on some of these outcomes is not surprising given PSNP transfers were not timely,

[13] Selected heterogeneous impact results are reported in Appendix C. The decisions around which heterogeneous impacts to examine and report are summarized in Appendix D, in the form of responses to reviewer comments.

unpredictable, and entitlements were incomplete (see chapter 6 of the Performance Report). These programme implementation challenges and their implications to outcomes need to be seen against the shocks experienced (discussed in section 5.1) and other broader macroeconomic challenges faced in the last three years.

### 5.5 Summary

This chapter has presented average trends in household assets, shocks, and forced sales of these assets to fulfill household food requirements and emergency cash needs and resilience to it. We have focused on livestock and productive assets in this chapter. On average, between baseline and end-line, livestock ownership of PSNP households has increased by about 0.5 TLU (compared to non-PSNP that grew by 0.2 TLU on average). The impact estimates also show that the PSNP has increased livestock TLU by 25 percent – this impact is statistically significant at 95 significance level. Moreover, the average value of livestock assets has also more than doubled in the two periods (although this change does not appear to be associated with the PSNP and show a counter intuitive result perhaps due to measurement error). Female headed households are less endowed with livestock assets than their male counterparts and that female headed PSNP households are even less endowed with livestock as compared to non-PSNP female headed households. I

Droughts remain the most important shocks among PSNP and non-PSNP households, followed by floods and erosion. Drought shock is reported highest among households in Amhara, followed by Oromia and SNNP. Overall, about 15 percent of PSNP (13 percent of non-PSNP) households from Amhara report drought as first most important shock. PSNP households are more likely to report experience of shocks than non-PSNP households. While the PSNP is designed to mitigate some of these shocks, the evidence from chapter 6 of the Performance Report shows that payments were neither predictable, nor made on timely fashion or in full entitlements.

In line with this, reports of distress asset sales are widespread. Increasing trends are observed over time in the proportion of households reporting distress sales of assets. The most reported distress asset sales are livestock assets for food and emergency cash needs. However, PSNP households are more likely to report these sales across all regions. Consistent with this descriptive narrative, results from the impact evaluation show that the PSNP4 did not have statistically significant impact on reducing distress asset sales, mainly on forced sales of livestock or productive assets for food or emergency cash needs.

# Chapter 6: Labor Allocation and Work Intensity

## 6.1 Introduction

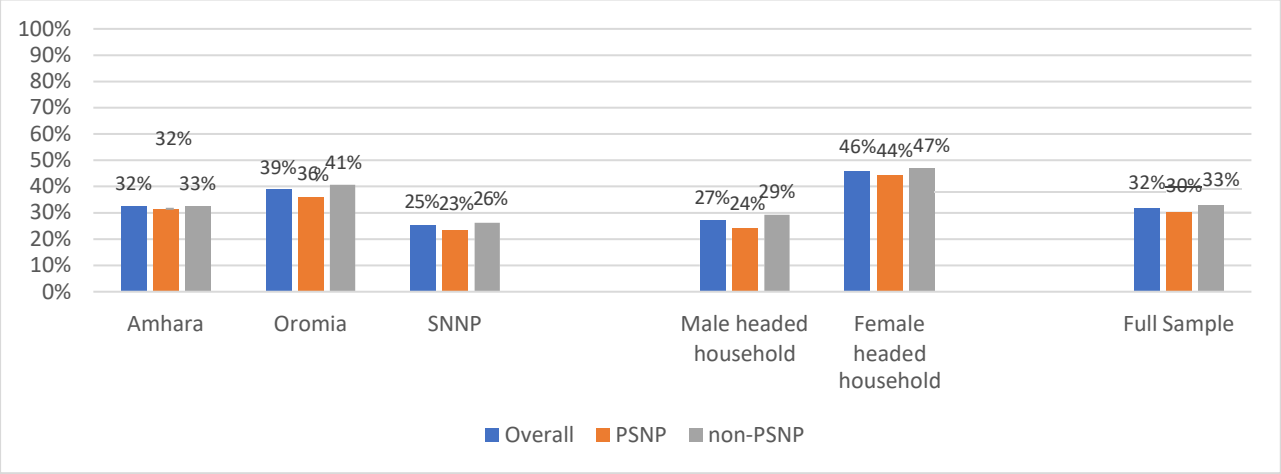
In this chapter, we present results on labor allocation and work intensity. We also provide information on diversification of income generation activities among households at end-line.

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

## 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. In this chapter we present information on labor allocation among household members 18 years and older. Overall, 32 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. There are some regional differences, as seen in Figure 6.1. We find that households in SNNP are least likely (25%) to report this and households in Oromia are most likely (39%). Non-PSNP households are more likely to report not having spent any time on these activities as compared to PNSP households, overall and by region. The difference in working worked or not is more pronounced when we compare male-headed and female-headed households (Figure 6.1), with 46 percent of female-headed households reporting that they spent zero hours on these activities in the past 7 days as compared to 27 percent among the male-headed households.

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

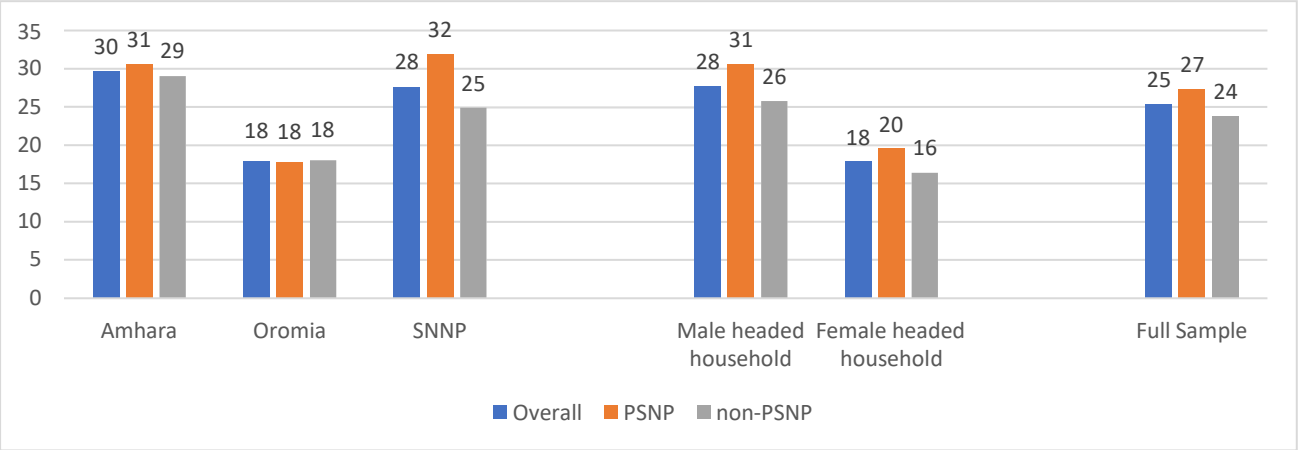



***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

Conditional on spending any time on labor activities (including agriculture, non-agriculture, salaried work or public works), the average total numbers of hours spent by the household in the 7 days prior to the survey is 25 hours. There is some variation in total hours spent on labor activities across regions with Amhara at 30 hours, Oromia at 18 hours and SNNP with an average of 28 hours (Figure 6.2). We

also observe that male-headed households spend 28 hours on average in the previous 7 days as compared to 18 hours among the female-headed households.

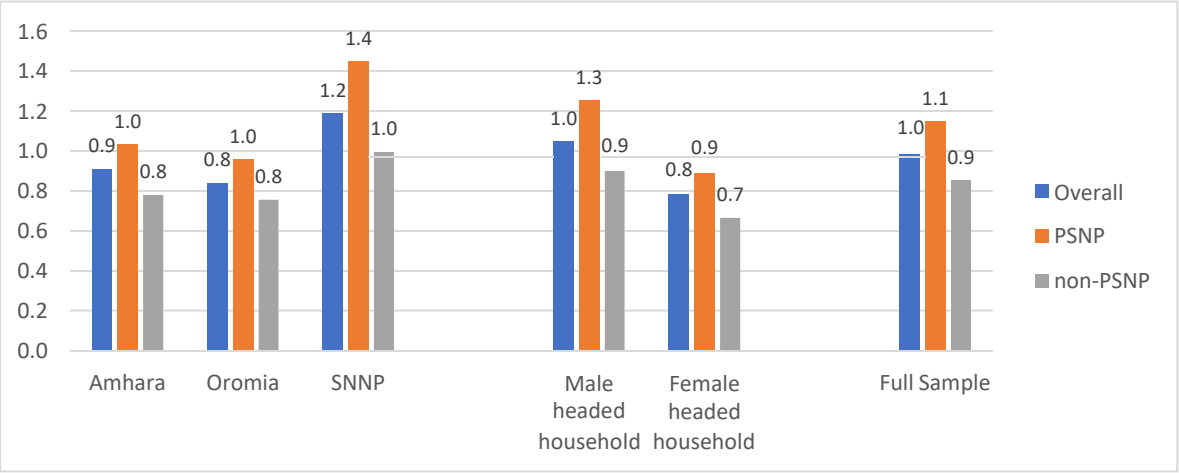
**Figure 6.2: Total hours spent by households on labor activities, by region, headship and PSNP status**




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

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 1 out of these 5 activities in the last 7 days. This is slightly higher at 1.1 among the PSNP households as compared to 0.9 among non-PSNP households with some variation across regions and sex of the household head (Figure 6.3).

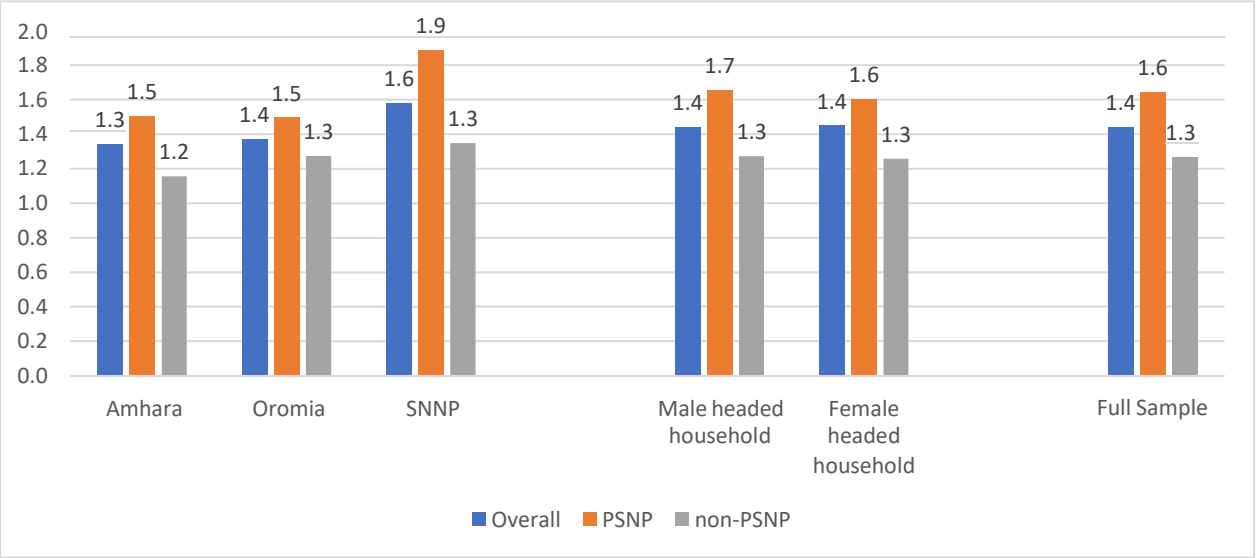
**Figure 6.3: Total number of activities involved in last 7 days, by region, headship and PNSP status**




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

When we restrict the sample to those who were engaged in any activities the last 7 days, these averages increase to 1.4 activities overall, 1.6 activities among the PSNP households and 1.3 activities among non-PSNP households (Figure 6.4).

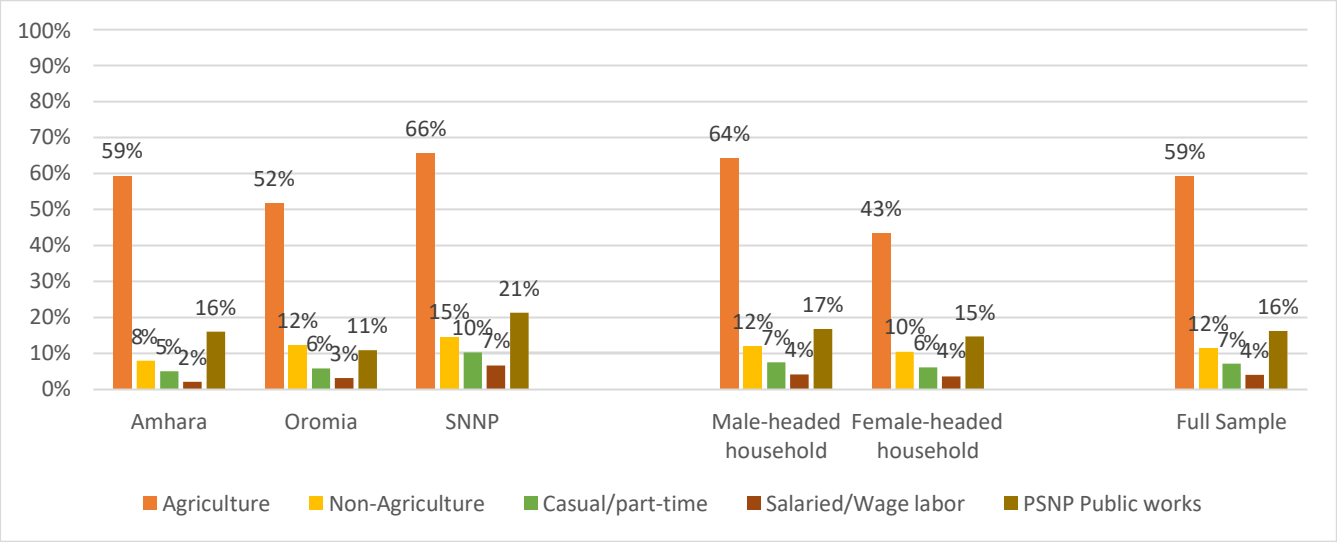
**Figure 6.4: Among those involved in any activity total number of activities involved in last 7 days, by region, headship and PNSP status**




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

On average 59% 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 is reported by 16% of the households and as expected this is concentrated among the PSNP households (35%, refer to Table 6.1).

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




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

We find that overall PSNP households are more likely to have worked in the 7 days prior to the survey, they are also engaged in a 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. In the next section, we investigate whether participating in the PSNP had an impact on these labor allocation and intensity outcomes.

We also construct an indicator of income diversification using information at the household level about whether the household was engaged in the following activities in the past 12 months prior to

the survey – wage work, casual agricultural wage work, off-farm business activities, received remittances. The average of income diversification outside of own agriculture is 0.17 among PSNP households and 0.18 among non-PSNP households. However, this difference is not statistically significant.

**Table 6.1: Labor outcome summary by the Treatment and Control samples in 2021**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean | N | Mean | N | |
| Total hours spent by household, 7 days | 27.360 | 2,234 | 23.811 | 2,688 | 0.000 |
| No. activities HH engaged, 7 days | 1.147 | 2,234 | 0.851 | 2,688 | 0.000 |
| HH engaged in ag activities in last 7 days | 0.573 | 2,234 | 0.611 | 2,688 | 0.006 |
| HH engaged in non-ag activities in last 7 days | 0.112 | 2,234 | 0.119 | 2,688 | 0.436 |
| HH engaged in casual labor in last 7 days | 0.070 | 2,234 | 0.072 | 2,688 | 0.789 |
| HH engaged in salaried work in last 7 days | 0.044 | 2,234 | 0.038 | 2,688 | 0.233 |
| HH engaged in PW in last 7 days | 0.348 | 2,234 | 0.012 | 2,688 | 0.000 |
| Percent hours HH engaged in ag activities last 7 days | 0.609 | 1,562 | 0.795 | 1,803 | 0.000 |
| Percent hours HH engaged in non-ag activities last 7 days | 0.077 | 1,562 | 0.111 | 1,803 | 0.000 |
| Percent hours HH engaged in casual labor last 7 days | 0.042 | 1,562 | 0.053 | 1,803 | 0.066 |
| Percent hours HH engaged in salaried work last 7 days | 0.027 | 1,562 | 0.032 | 1,803 | 0.280 |
| Percent hours HH engaged in PW last 7 days | 0.246 | 1,562 | 0.009 | 1,803 | 0.000 |
| Income diversification outside of own agriculture (0-4) | 0.176 | 2,235 | 0.184 | 2,691 | 0.527 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the PSNP and non-PSNP samples.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

### 6.3 Impact of the PNSP on labor allocation and work intensity

Table 6.2 presents the impact on labor allocation and work intensity outcomes. We find that PSNP had a positive and significant impact on the total number of hours spent on work activities by the household in the 7 days prior to the survey. PSNP household spent 5.9hours more than the comparison mean of 24 hours. We also observe a positive and statistically significant impact on total number of activities the household was engaged in, in the previous 7 days – 0.4 activities which is about 47 percent of the comparison mean. We do not observe an impact on the probability of being engaged in any specific activity except for small positive impact on the likelihood of being engaged in salaried work and large positive impact on likelihood of being engaged in PSNP public works in the last 7 days.

Among those that were engaged in some labor activity in the last 7 days, there seems to be some reallocation of time across activities. We find that the PSNP had a negative impact on the proportion

of hours spent on agricultural and non-agricultural activities and a positive impact on proportion of time spent on the public works. PSNP decreased the proportion of time spent on agricultural activities by 18 percentage points in the last 7 days and increased the proportion of time spent on PWs by 24 percentage points. We find that PSNP has no impact on income diversification outside of own agricultural activities.

**Table 6.2: Impact on labor allocation and work intensity, Nearest Neighbor Matching**

| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| Total hours spent by household, 7 days | 5.95*** | 4,350 | 23.811 |
| | (1.017) | | |
| No. activities HH engaged, 7 days | 0.397*** | 4,350 | 0.851 |
| | (0.027) | | |
| HH engaged in ag activities in last 7 days | -0.009 | 4,350 | 0.611 |
| | (0.014) | | |
| HH engaged in non-ag activities in last 7 days | 0.012 | 4,350 | 0.119 |
| | (0.010) | | |
| HH engaged in casual labor in last 7 days | 0.014 | 4,350 | 0.072 |
| | (0.008) | | |
| HH engaged in salaried work in last 7 days | 0.019*** | 4,350 | 0.038 |
| | (0.006) | | |
| HH engaged in PW in last 7 days | 0.361*** | 4,350 | 0.012 |
| | (0.011) | | |
| Percent hours HH engaged in ag activities last 7 days | -0.213*** | 3,454 | 0.795 |
| | (0.014) | | |
| Percent hours HH engaged in non-ag activities last 7 days | -0.023*** | 2,997 | 0.111 |
| | (0.008) | | |
| Percent hours HH engaged in casual labor last 7 days | -0.000 | 2,997 | 0.053 |
| | (0.006) | | |
| Percent hours HH engaged in salaried work last 7 days | -0.001 | 2,997 | 0.032 |
| | (0.005) | | |
| Percent hours HH engaged in PW last 7 days | 0.236*** | 2,997 | 0.009 |
| | (0.010) | | |
| Change in income diversification, baseline to end-line | 0.022 | 4,924 | -0.090 |
| | (0.020) | | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

## 6.4 Summary

In this chapter we presented descriptive and impact results on labor allocation and work intensity in the 7 days prior to survey among households members 18 years and older. 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.

- 32 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 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 1 out of 5 activities in the last 7 days. This is slightly higher at 1.1 among the PSNP households as compared to 0.9 among non-PSNP households with some variation across regions and sex of the household head.
- On average 59% 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 is reported by 16% of the households and as expected this is concentrated among the PSNP households (35%, refer to Table 6.1).

We set out to answer the research question, **RQ11: Has participation in the PSNP influenced the labor allocation and work intensity decisions of beneficiary households?**

- The impact analysis shows that participation in the PSNP had a positive and statistically significant impact on total hours spent in the last 7 days by the household.
- PSNP also had a positive and statistically significant impact on total number of activities the household was engaged in, in the previous 7 days – 0.4 activities which is about 47 percent of the comparison mean.

# Chapter 7: Poverty

## 7.1 Introduction

In this chapter, we present results on poverty using objective as well as 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 PSNP on these poverty related outcomes. This chapter will address **RQ3: Has PSNP reduced poverty?**

In addition to poverty measures, we also provide descriptive results on indicators for agency and intra-household decision making.

## 7.2 Consumption poverty

We measure poverty among households using the national food poverty line and the consumption expenditure data collected at during the end-line survey. In principle, there are two ways of accommodating temporal price changes in poverty analysis. The first is to measure expenditures in all rounds in current prices and adjust the poverty line for inflation that occurred between rounds. The other is to measure expenditures in all rounds in constant prices of a given reference round or year (i.e., adjust the expenditures for inflation that occurred between rounds) and use the poverty line expressed in the reference year prices for all rounds/years. We adopted the second approach, primarily because the price data we have for 2021 covers a more limited list of commodities. We use the 2016 national poverty line with household consumption expenditures in both rounds (baseline and end-line) measured using 2016 prices.

As seen in Table 7.1, over 58% of the PSNP households are categorized as poor based on the national food poverty line (315 Ethiopian Birr per adult equivalent per month in 2016 prices) and 55% of non-PSNP households are categorized as poor based on the same national food poverty line. This difference is statistically significant (p=0.03).

**Table 7.1: Poverty status summary by the Treatment and Control samples in 2021**

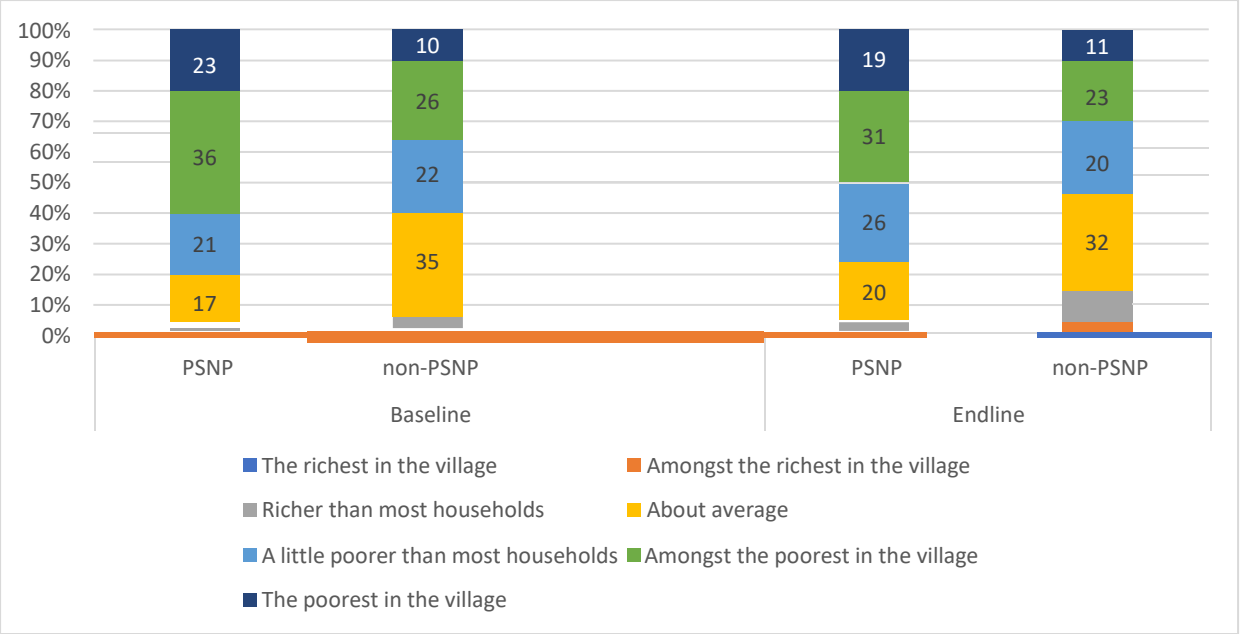
| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Households defined as poor based on national food poverty line[a] (%) | 57.7 | 1.04 | 2,223 | 54.6 | 1.04 | 2,291 | 0.03 | 0.116 |

***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. [a] The national food poverty line is Birr315 per adult equivalent per month expressed in 2016 prices.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

### 7.3 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 self-perception of economic standing is lower than the non-PSNP households at baseline and at end-line. We present the summary statistics by PSNP status and survey round in Table 7.2. We find that at baseline there is a statistically significant difference between the proportion of households reporting all levels of economic standing in the village, except for being the richest in the village and little poorer than most households in the village. PSNP households are more likely than non-PSNP households to report themselves as poorer than others at baseline. At end-line, the difference in reported economic standing between PSNP and non-PSNP households is statistically significant for all levels reported and PSNP households are more likely than non-PSNP households to perceive themselves as poorer than others in the village. Nevertheless, the fraction of PSNP households that identify themselves as the poorest or among the poorest in the village compared to their neighbors have fallen between baseline and end-line.

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




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

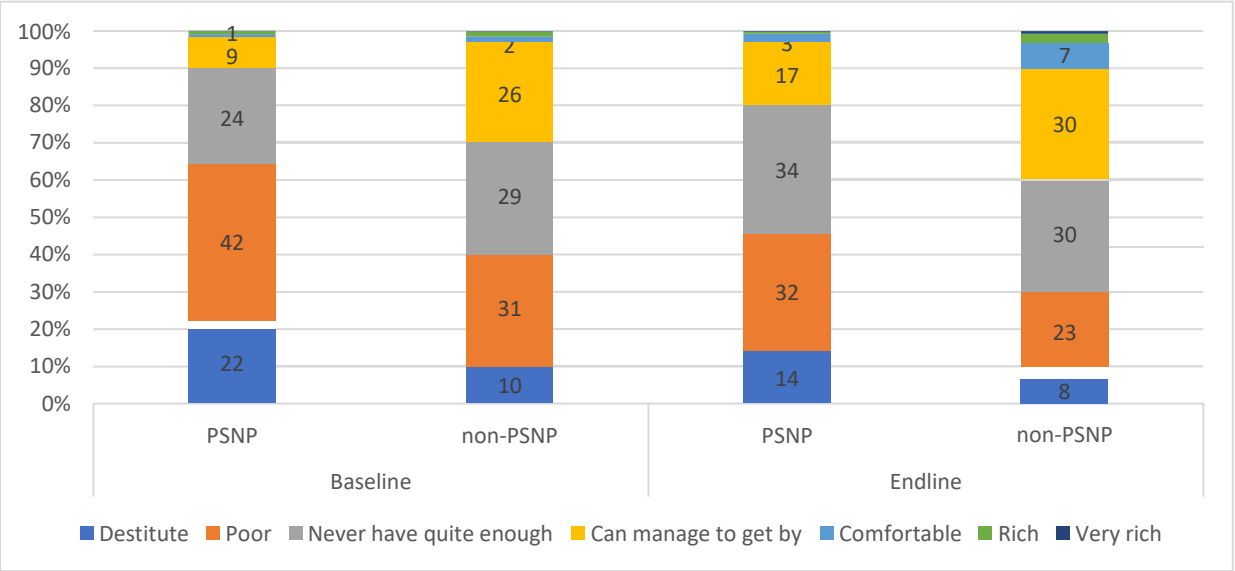
**Table 7.2: Self-perceived poverty summary by the Treatment and Control samples by survey round**

| | Baseline | | | | | End-line | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSNP | | Non-PSNP | | P-value of difference by PNSP status | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
| | Mean (%) | N | Mean (%) | N | | Mean (%) | N | Mean (%) | N | |
| The richest in the village | 0 | 2,233 | 0 | 2,689 | 0.046 | 1 | 2,235 | 1 | 2,690 | 0.006 |
| Amongst the richest in the village | 2 | 2,233 | 2 | 2,689 | 0.113 | 1 | 2,235 | 3 | 2,690 | 0.000 |
| Richer than most households | 1 | 2,233 | 5 | 2,689 | 0.000 | 3 | 2,235 | 10 | 2,690 | 0.000 |
| About average | 17 | 2,233 | 35 | 2,689 | 0.000 | 20 | 2,235 | 32 | 2,690 | 0.000 |
| A little poorer than most households | 21 | 2,233 | 22 | 2,689 | 0.400 | 26 | 2,235 | 20 | 2,690 | 0.000 |
| Amongst the poorest in the | 36 | 2,233 | 26 | 2,689 | 0.000 | 31 | 2,235 | 23 | 2,690 | 0.000 |
| The poorest in the village | 23 | 2,233 | 10 | 2,689 | 0.000 | 19 | 2,235 | 11 | 2,690 | 0.000 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

We also asked respondents to describe their household focusing solely on their own household's circumstances. The options included were very rich, rich, comfortable, can manage to get by, never have quite enough, poor, and destitute. We consider this as their perception of absolute poverty which we refer to as subjective poverty. This is presented in Figure 7.2 by PSNP status and survey round. We observe that, at baseline, a large percentage of PSNP households (64%) consider themselves as destitute or poor. The corresponding percentage among non-PSNP households is lowerat 41%. There is also about a quarter of non-PSNP household that think they "can manage to get by" whereas only 9% of the PSNP households reported this. We see some improvements over time in bothgroups. Among the PSNP households, the group of destitute and poor households reduced to 14% and 36%, respectively, and reduced to 8% and 23%, respectively, among the non-PSNP households. The differences in these means are statistically significant, as reported in Table 7.3.

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



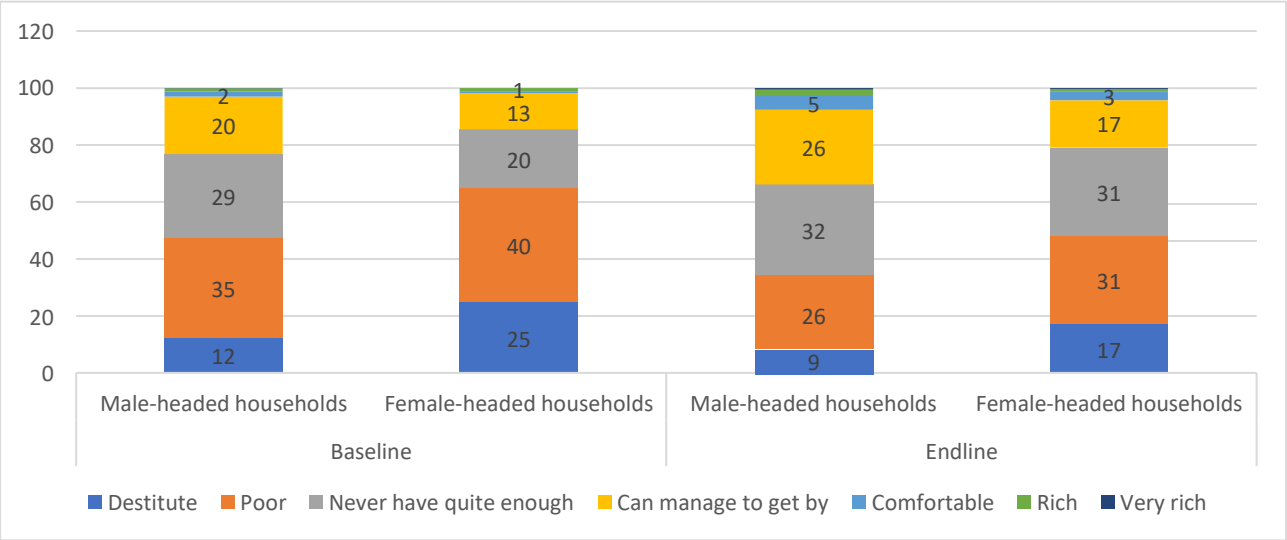

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

**Table 7.3: Subjective poverty summary by the Treatment and Control samples by survey round**

| | Baseline | | | | | End-line | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PSNP | | Non-PSNP | | P-value of difference by PNSP status | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
| | Mean (%) | N | Mean (%) | N | | Mean (%) | N | Mean (%) | N | |
| Destitute | 22 | 2,231 | 10 | 2,684 | 0.000 | 14 | 2,235 | 8 | 2,690 | 0.000 |
| Poor | 42 | 2,231 | 31 | 2,684 | 0.000 | 32 | 2,235 | 23 | 2,690 | 0.000 |
| Never have quite enough | 24 | 2,231 | 29 | 2,684 | 0.000 | 34 | 2,235 | 30 | 2,690 | 0.001 |
| Can manage to get by | 9 | 2,231 | 26 | 2,684 | 0.000 | 17 | 2,235 | 30 | 2,690 | 0.000 |
| Comfortable | 1 | 2,231 | 2 | 2,684 | 0.000 | 3 | 2,235 | 7 | 2,690 | 0.000 |
| Rich | 1 | 2,231 | 1 | 2,684 | 0.104 | 1 | 2,235 | 3 | 2,690 | 0.000 |
| Very rich | 22 | 2,231 | 10 | 2,684 | 0.000 | 14 | 2,235 | 8 | 2,690 | 0.000 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

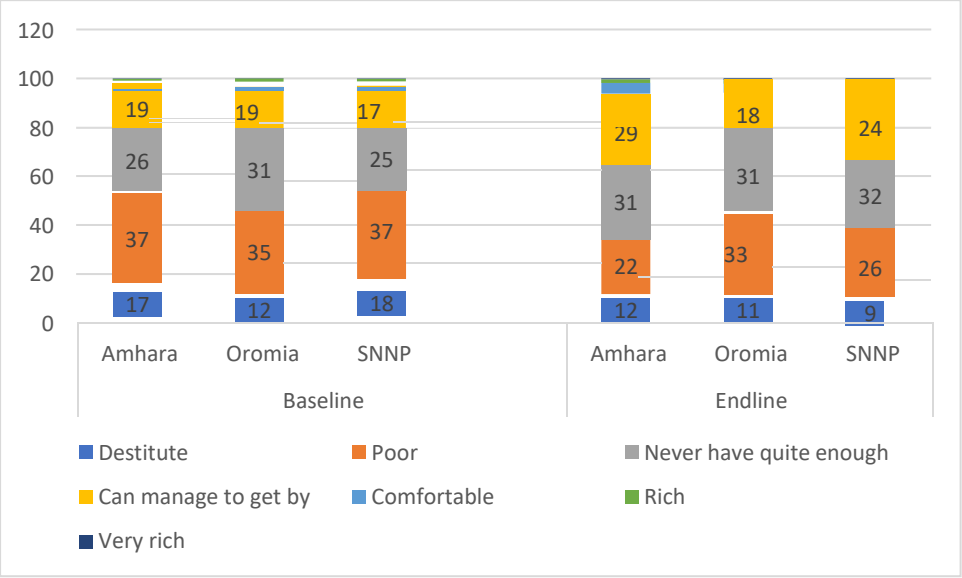
**Figure 7.3: Subjective poverty, by headship and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

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 improvement. The three regions had very similar levels of subjective poverty at baseline. Over time, Amhara and SNNP have seen some improvements.

**Figure 7.4: Subjective poverty, by region and survey round**

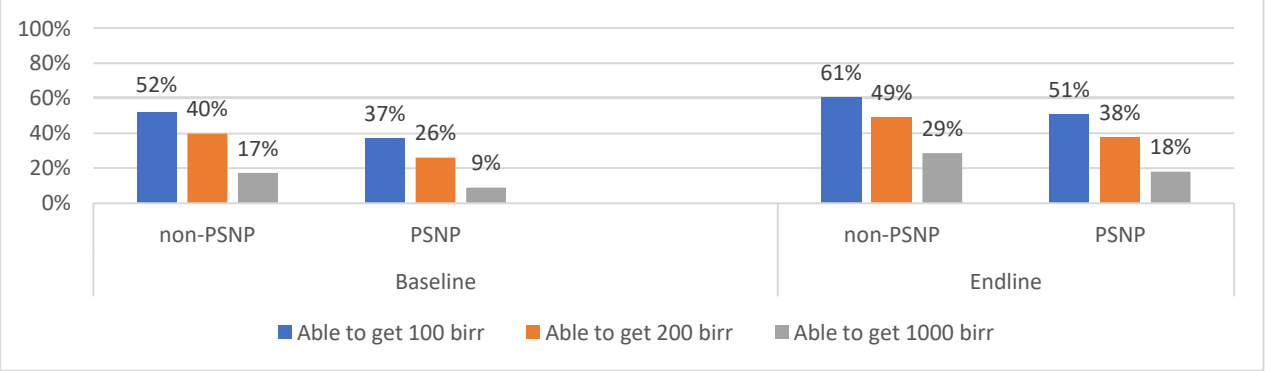



***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

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. 37% of the PNSP households

and 52% of non-PSNP households, at baseline, were able to obtain 100 Birr. As the amount increases the proportion that can obtain it drops. Over time, we observed a considerable improvement in the PSNP households' ability to obtain such emergency funds.

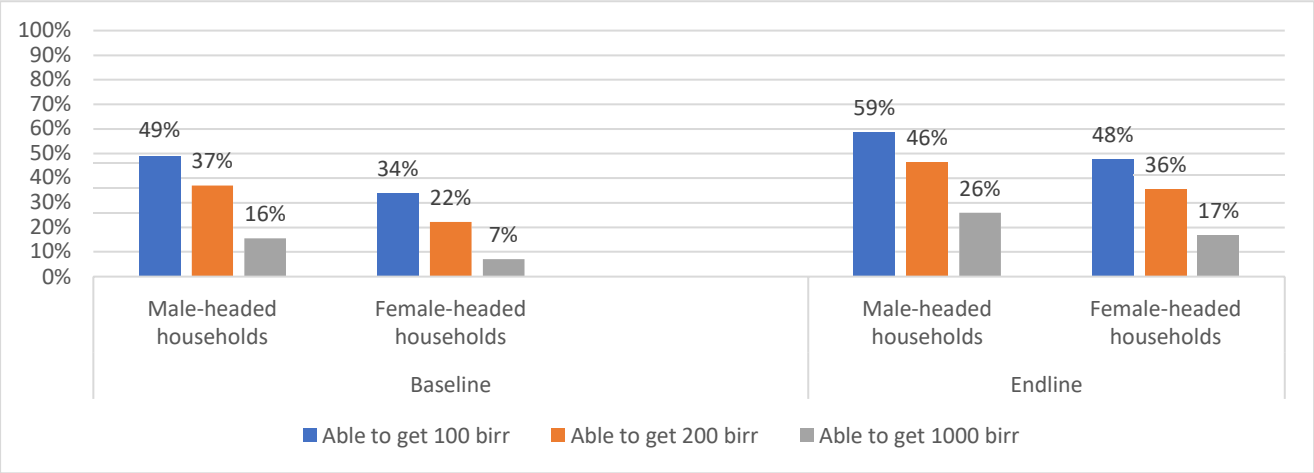
**Figure 7.5: Able to obtain emergency funds, by PSNP status and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

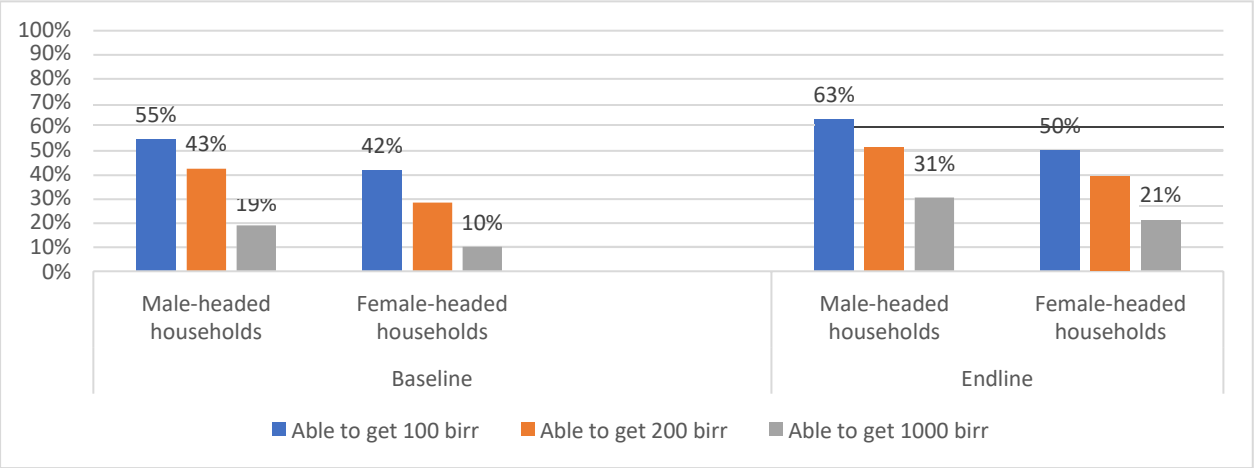
Comparing male- and female-headed households (Table 7.6a), we find that a larger percentage of male-headed households are able to obtain emergency funds at baseline but, by end-line, both groups have seen improvements in their ability to raise emergency funds. The pattern is similar when we disaggregate this further by PSNP status as seen in Figures 7.6b and 7.6c.

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



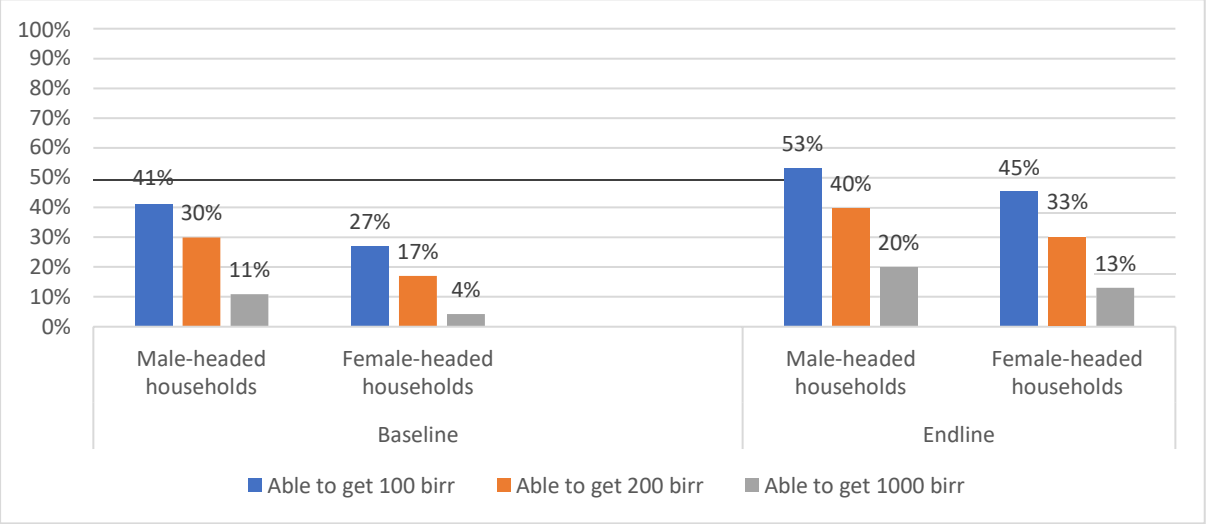

***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

**Figure 7.6b: Able to obtain emergency funds among non-PSNP households, by headship and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

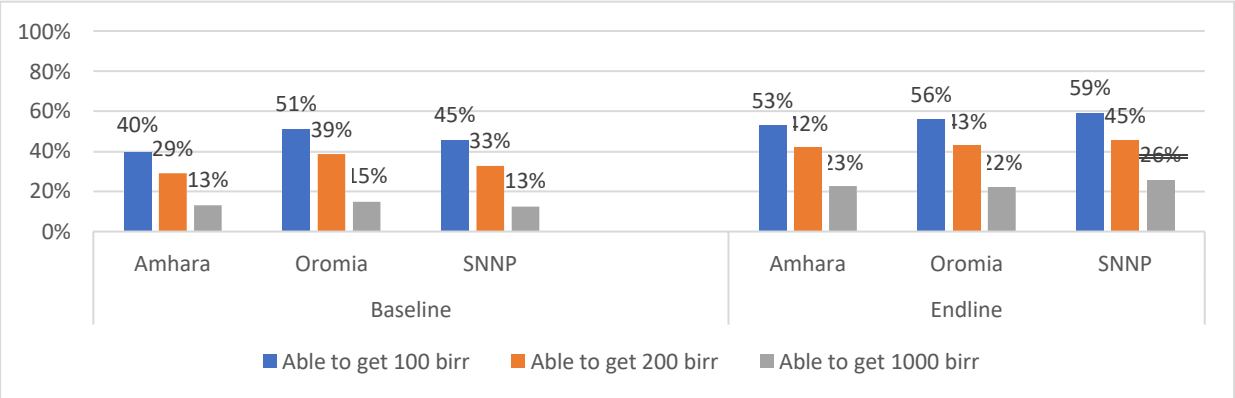
**Figure 7.6c: Able to obtain emergency funds among PSNP households, by headship and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

There is some regional variation in the ability to raise emergency funds with the largest percentage of households in Oromia able to do so (Figure 7.7). All regions observed an improvement.

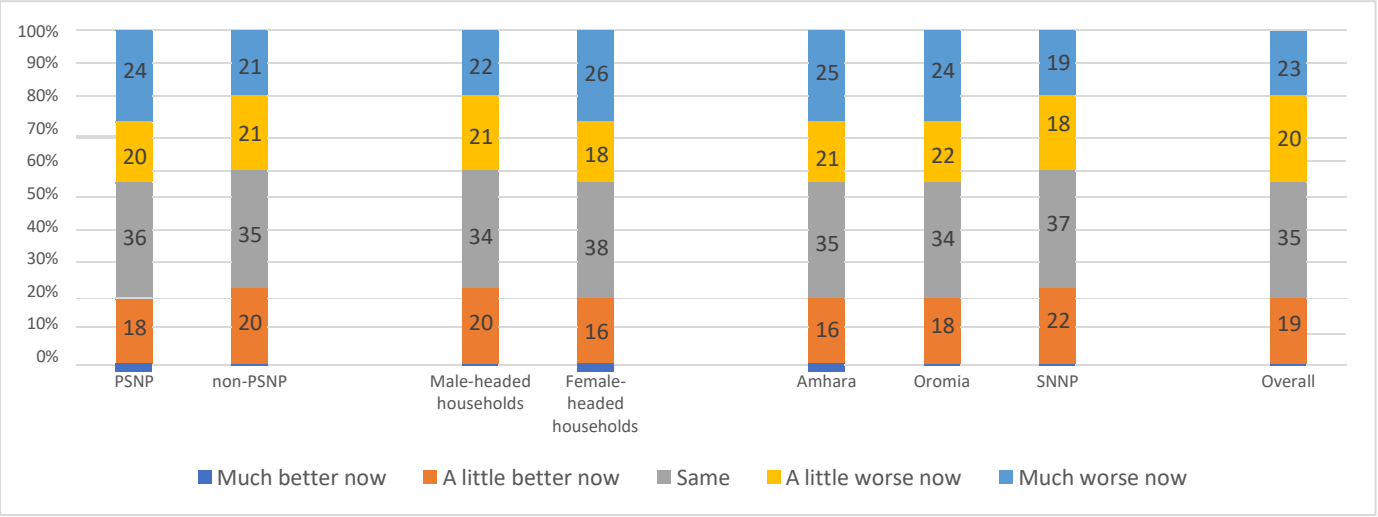
**Figure 7.7: Able to obtain emergency funds, by region and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

At end-line, we also asked households about their circumstances two years ago (Figure 7.8). 23% of the households overall reported doing much worse now, 20% reported little worse, 35% reported doing same as two years ago and 19% reported doing a little better now. We do not observe significant variations across different groups. While we observe some improvements in relative economic standing, subjective poverty and 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.

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




***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Surveys.

## 7.4 Impact of the PSNP on poverty

### *7.4.1 Impact on consumption poverty*

Table 7.4 presents the impact estimates on the food poverty indicator and we find that the point estimate is positive but not statistically significant. This implies that PNSP had no impact on the likelihood of being classified as poor based on the national food poverty line.

**Table 7.4: Impact results from NNM - Outcome group 8 - Poor National**

| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| % Households defined as poor based on national food poverty line[a] | 0.024 | 4,508 | 0.55 |
| | (0.016) | | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$. [a] National food poverty line is Birr 315 per adult equivalent per month in 2016 prices)

### *7.4.2 Impact 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 7.5 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. This result remains statistically insignificant in the 95% trimmed sample model (See Table 7.7).

**Table 7.5: Impact on subjective poverty, Nearest Neighbor Matching**

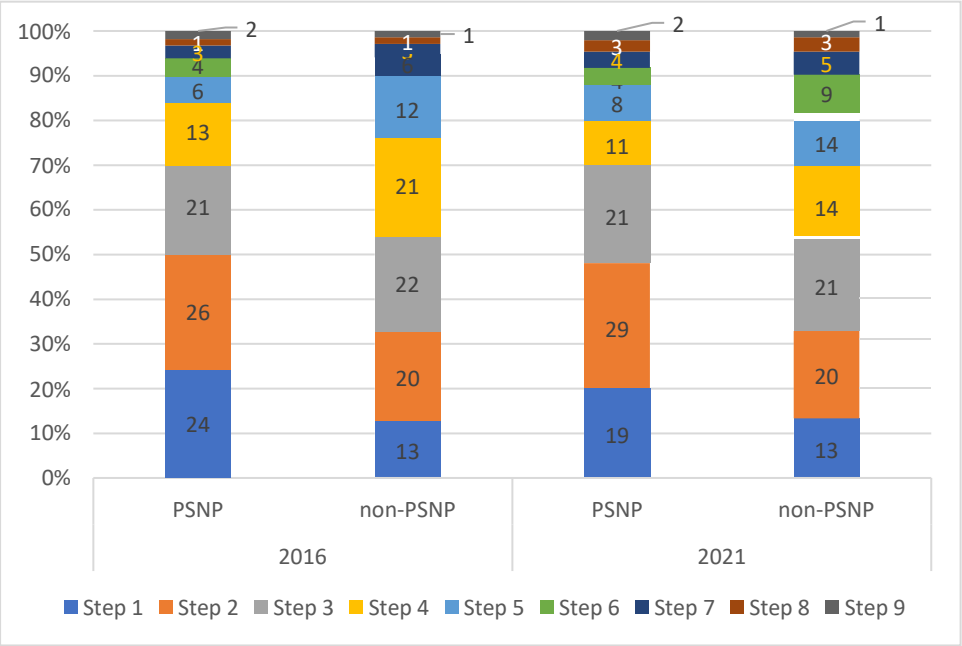
| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| Change in subjective poverty | 0.014 | 4,913 | 0.317 |
| | (0.040) | | |

Notes: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

## 7.5 Agency and intrahousehold decision making

In this section, we present descriptive results on variables that reflect individuals' agency, comfort in speaking in public and intrahousehold decision making. We begin with a question, administered at baseline and end-line, which tells households to imagine again 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 households are more likely to be on lower steps on the ladder than non-PSNP households and there has been minimal improvement over time.

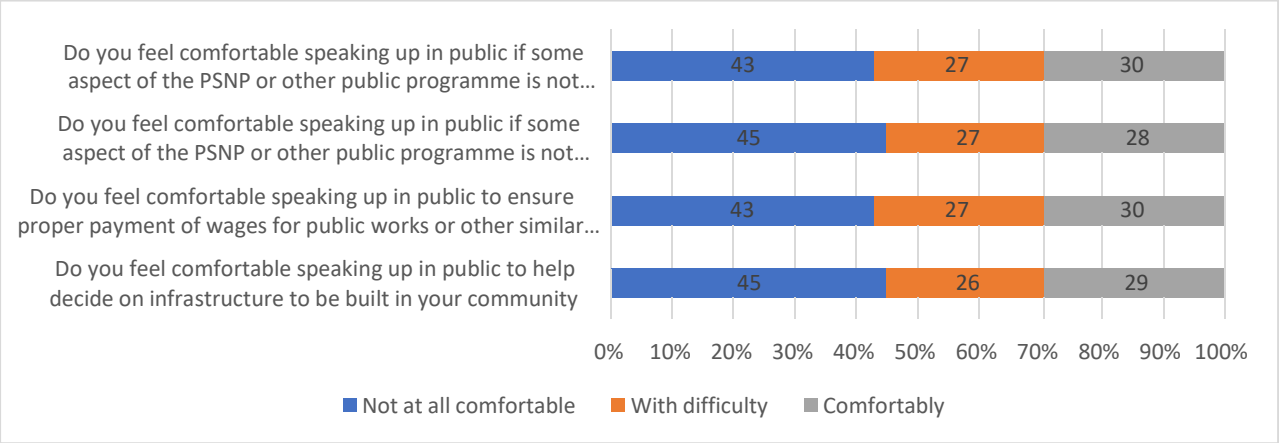
**Figure 7.9: Control over life – 9-step ladder, by PSNP status and survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

Figure 7.10 shows descriptive results, for all households, 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 more than 40% of the households are not at all comfortable speaking up in public. About 30% are comfortable and the remaining 26-27% would speak up with difficulty. This reflects the head's comfort in speaking publicly.

**Figure 7.10: Comfort in public speaking at 2021**

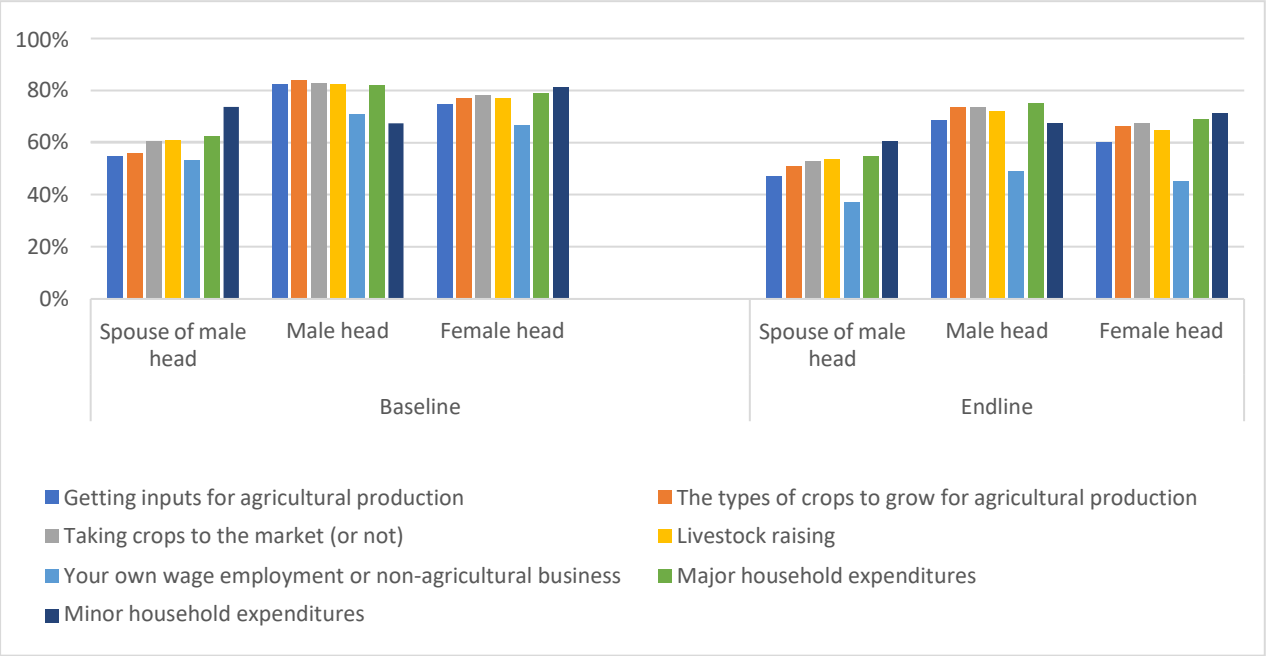



***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

We next examine intrahousehold decision making by comparing "say in decision making" – on matters relating to agriculture, livestock, employment and household expenditures – across the male head, spouse of the male head and the female head. Figure 7.11 summarises the responses from household heads (male and female) and spouse of male head for whether they have a say in decision-making around different matters. We observe that, overall, the male heads are most likely to be able to make

their own decisions as compared to the spouses of male heads and their female counterparts. Over time, this ability to make a decision on their own has decreased across all groups.

**Figure 7.11: Intrahousehold decision making, by survey round**




***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

## 7.6 Summary

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

- Starting with relative subjective wellbeing, PSNP households' perception of their economic standing is lower than the non-PSNP households' at both baseline and end-line. We do not observe much movement in this perception over time.
- In terms of perception of absolute poverty, at baseline, a large percentage of PSNP households (64%) consider themselves as destitute or poor. The corresponding percentage among non-PSNP households is lower at 41%. There is also about a quarter of non-PSNP household that think they "can manage to get by" whereas only 9% of the PSNP households reported this. We see some improvements over time in both groups. Among the PSNP households, the group of destitute and poor households reduced to 46% and it reduced to 31% 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 improvement. The three regions had very similar levels of subjective poverty at baseline. Over time, Amhara and SNNP have seen some improvements.

- 37% of the PNSP households and 52% of non-PSNP households, at baseline, reported that would be able to obtain 100 Birr for an emergency. As the amount increases the proportion that can obtain it drops. Over time, we observed a considerable improvement in the PSNP households' ability to obtain such emergency funds.
- 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.
- 23% of the households overall reported doing much worse now as compared to two years ago, 20% reported little worse, 35% reported doing same as two years ago and 19% reported doing a little better now. 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 observe that PSNP had no impact on the likelihood of being poor based on the national food poverty line.
- We do not find any impact of the PSNP on subjective poverty

This chapter also presents descriptive on agency and intrahousehold decision making.

- PSNP households are less likely to feel they have full control over their life as compared to non-PSNP households. And there has been minimal improvement over time.
- 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 households are not at all comfortable, about 30% are comfortable and the remaining 26-27% would speak up with difficulty.
- In terms of "say in decision making" – on matters relating to agriculture, livestock, employment and household expenditure – overall the male heads are most likely to be able to make their own decisions as compared to the spouses of male heads and the female-heads.

**Table 7.6: Mean subjective poverty, by PSNP status in 2021**

| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Subjective poverty | 2.647 | 1.052 | 2,235 | 3.157 | 1.199 | 2,690 | 0.000 | -0.452 |

***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 data from the PSNP IV End-line (2021) Household Survey.

**Table 7.7: Impact results of 95% sample from NNM - Outcome group 3 - Poverty**

| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| Change in subjective poverty | 0.011 | 4,662 | 0.325 |
| | (0.041) | | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$
***Source***: Authors' computation using data from the PSNP IV Baseline (2016) and End-line (2021) Household Surveys.

# Chapter 8: Human Development Outcomes

## 8.1 Introduction

The fourth phase of the PSNP places a new emphasis on nutrition-sensitivity, with added interventions focusing on behavioral change and care practices. A particular focus is on pregnant and lactating women who are temporarily moved from the public works to direct support with health and nutrition-related soft conditionalities (referred to as co-responsibilities). Furthermore, public works clients are encouraged to attend nutrition behavioral change communication (BCC) sessions by considering the sessions as a part of their public works labor requirement.

This chapter provides descriptive results to understand RQ4: Has PSNP improved human development outcomes? It includes measures relating to pregnant women and access to antenatal care, maternal knowledge of IYCF practices[14], knowledge or and interactions with the health extension worker and the health development army, child labor. In addition, we also present impact estimates on mid-upper arm circumference among children and mothers. Since the samples are small and set of households do not overlap over rounds, we do not present changes over time because they are not feasible. We also present the statistics in tables which clearly present the sample size for each indicator. This is not always feasible to do in graphs and thus they can be misleading.

## 8.2 Pregnant women, PSNP, and antenatal care seeking

In this subsection we present summary statistics on pregnancy, experience with PSNP in the last pregnancy and care seeking in the last pregnancy. We start by examining the percentage of women in our end-line survey sample that have ever been pregnant since May 2016. As seen in Table 8.1, about 30% of the women were ever pregnant since May 2016 and about 10% of these women were pregnant at the time of the end-line survey. 23% of the households, where a woman had ever been pregnant since May 2016, were PSNP beneficiaries when they learned of the pregnancy. Among the PSNP beneficiary households that had a female household member that had ever been since May 2016,43% of these women were working on PW when they learned about the pregnancy. Among the PSNP beneficiary households that had a female household member that had ever been pregnant since May 2016, about 60% of the women also reported that someone else from their household was working on PW when they learned about the pregnancy. 63% of the women who were working of the PW when they found out about their pregnancy stopped working. 45% reported that a health worker advised them to stop working and 38% reported that the development agent advised them to stop working. Among the women who did not stop working about 18% continued to work until they gave birth. On average, women stopped working when they were 4 months pregnant and about 45% wished they could have stopped earlier. About 30% reported receiving direct support benefits after the women stopped working. And about 24% reported that other household members worked more to compensate for the woman not working. 64% report returning to PW after giving birth. On average

[14] We examine and present results on knowledge of IYCF and not the IYCF practices.

women went back after 7 months of giving birth and 68% of these women said that going back to work affected their ability to breastfeed their child.

**Table 8.1: Pregnancy and PSNP, end-line**

| | All | | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | Mean[1] | N | |
| Ever pregnant since May 2016 | 30 | 5,069 | 28 | 2,216 | 31 | 2,675 | 0.051 |
| Currently pregnant | 10 | 1,499 | 9 | 629 | 12 | 828 | 0.044 |
| Household receiving PSNP benefits when first learned about pregnancy | 23 | 1,499 | 50 | 629 | 4 | 828 | 0 |
| Woman working on public works when learned about pregnancy | 43 | 349 | 45 | 313 | 33 | 30 | 0.231 |
| Any other household members working on public works when first learned about pregnancy | 60 | 349 | 65 | 313 | 17 | 30 | 0 |
| Did you stop working on public works when you first learned you were pregnant | 64 | 210 | 63 | 204 | 100 | 5 | _ |
| Advised by a health worker to stop work | 45 | 210 | 44 | 204 | 80 | 5 | - |
| Advised by a Development Agent to stop work | 38 | 210 | 37 | 204 | 60 | 5 | - |
| Continued to work until gave birth | 18 | 199 | 17 | 195 | 33 | 3 | - |
| Number of months pregnant were stopped working[2] | 3.9 | 175 | 3.9 | 170 | 4.5 | 4 | - |
| Wished they could have stopped working earlier | 45 | 175 | 44 | 170 | 75 | 4 | - |
| Discouraged from stopping work by anyone on the Community Food Security Task Force | 28 | 210 | 28 | 204 | 20 | 5 | - |
| Household received any Direct Support payments after woman stopped working | 30 | 210 | 29 | 204 | 20 | 5 | - |
| Other household members worked more days to make up for woman not working | 24 | 210 | 25 | 204 | 20 | 5 | - |
| Resumed working on public works after giving birth | 64 | 210 | 63 | 204 | 0 | 3 | - |
| Number of months after giving birth woman resumed working[2] | 6.9 | 92 | 6.9 | 92 | | 0 | |
| Resuming working affect their ability to breastfeed their child | 63 | 92 | 63 | 92 | | 0 | |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise. [2] Number.

***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

We also asked about antenatal care received by the women during their last pregnancy (Table 8.2). About 26% of the women were pregnant at least once in the last 3 years. Among those that were pregnant in the last 3 years, 30% were visited by a health worker/community worker and about 49% reported receiving antenatal care during that pregnancy among PSNP households. About 31% reported being pregnant during the COVID period and about 21% reported receiving antenatal care during the pregnancy. Among PSNP households, the women were 4.8 months pregnant when they first visited the health facility, and they visited an average of 3.2 times during their pregnancy. The

non-PSNP households visited 3.7 times and this difference is statistically significant. 43% received counseling about nutrition of pregnant women and 45% received counseling about breastfeeding during their last pregnancy. We do not observe statistically significant differences between PSNP and non-PSNP households, except that women among non-PSNP households are more likely to have been pregnant in the last 3 years and they also visited the health center more often, these differences are statistically significant.

**Table 8.2: Antenatal care, end-line**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Pregnant any time in the last 3 years | 25 | 2,216 | 28 | 2,675 | 0.015 |
| When pregnant with last child, any health worker or community worker visit home | 30 | 555 | 28 | 753 | 0.435 |
| Received antenatal care during this pregnancy | 49 | 555 | 52 | 753 | 0.413 |
| Pregnant during the COVID-19 period | 31 | 555 | 32 | 753 | 0.759 |
| Received antenatal care during this (COVID-19 period) pregnancy | 21 | 555 | 22 | 753 | 0.619 |
| During the last pregnancy, number of months pregnant when first visited health facility[2] | 4.8 | 116 | 4.7 | 166 | 0.582 |
| Number of times visited the health facility (for antenatal care) during last pregnancy[2] | 3.2 | 116 | 3.7 | 166 | 0.036 |
| During the last pregnancy, received counseling or information about nutrition for pregnant women | 43 | 555 | 44 | 753 | 0.642 |
| During the last pregnancy, received counseling about breastfeeding | 45 | 555 | 46 | 753 | 0.856 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise. [2] Number.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

## 8.3 Maternal IYCF knowledge

In the end-line survey, we also asked all mothers in the sample of their knowledge of breastfeeding and complementary feeding, shown in Table 8.3. On average, 65% of the women in PNSP households knew that children should be breastfed immediately after birth and 61% knew that a child should be exclusively breastfed for the first 6 months. Knowledge of timely introduction of complementary foods is worse. We observe that about 41% and 34% of the women in PSNP households know that liquids and solid foods, respectively, can be introduced at 6 months of age. The only difference between PSNP and non-PSNP mothers is in breastfeeding knowledge. 65% of PSNP mothers know about immediate initiation of breastfeeding as compared to 62% of non-PSNP mothers. And 61% of

PSNP mothers know about exclusive breastfeeding till 6 months as compared to 58% of non-PSNP mothers. Although these differences are statistically significant, the absolute differences are not large.

**Table 8.3: IYCF Knowledge among mothers, end-line**

| Mothers know that… | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Child should be breastfed immediately after birth | 65 | 2,465 | 62 | 3,013 | 0.083 |
| Child should be exclusively breastfed until 6 months of age | 61 | 2,465 | 58 | 3,013 | 0.022 |
| Liquids should be introduced at 6 months of age | 41 | 2,465 | 41 | 3,013 | 0.875 |
| Solids should be introduced at 6 months of age | 34 | 2,465 | 35 | 3,013 | 0.494 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

## 8.4 Child health and nutrition outcomes

For all children under the age of 5 years, data was collected on health monitoring. As seen in Table 8.4, about 56% and 58% of the children in PSNP and non-PSNP households, respectively, had a health card. However, a much smaller percentage of children were measured. About 17%, 16% and 20% were measured for weight, height and mid-upper arm circumference among PSNP households, respectively. About 24% received advice or information about child feeding and only 3% of the children were identified as severely malnourished among PSNP households. It appears that children among PSNP households are slightly more likely to be measured as compared to those that belong to non-PSNP households (statistically significant as indicated in Table 8.4).

**Table 8.4: Summary statistics of Child health**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Child has a health card | 56 | 799 | 58 | 1,039 | 0.54 |
| Weight was measured in the past 3 months | 17 | 799 | 15 | 1,039 | 0.244 |
| Height was measured in the past 3 months | 16 | 799 | 12 | 1,039 | 0.01 |
| Mid-upper arm circumference was measured in past 3 months | 20 | 799 | 16 | 1,039 | 0.02 |
| Received advice or information about feeding the child | 24 | 799 | 20 | 1,039 | 0.014 |
| Child was identified as being severely malnourished in the past 6 months | 3 | 799 | 3 | 1,039 | 0.326 |
| Received any specific food or milk as treatment for severe acute malnutrition | 56 | 27 | 63 | 27 | 0.588 |
| Received a referral to a facility to receive treatment for severe malnutrition | 11 | 27 | 11 | 27 | 1.000 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

## 8.5 Health extension worker and health development army

PSNP households were more likely than non-PSNP households to know the health extension worker (HEW) (42% vs 38%, p<0.01). Among those who know the HEW, – PSNP households are more likely to report that were visited by the HEW at home in the past 3 months as compared to non-PSNP households (31% vs 26%, p<0.05). During the last home visit, 65% of the PSNP households reported that the HEW talked about breastfeeding, child feeding and nutrition. And among those households who know the HEW, 28% of the PSNP households had contact with them in the community (outside their home or the health post) whereas 23% of non-PSNP households reported having contact with them in the community and this difference is statistically significant. Compared to pre-COVID times, 64% of the PSNP respondents reported that the frequency of HEW home visit increased31% of the PSNP respondents reported visiting the health post in the last 3 months.

**Table 8.5: Interactions with health extension workers (HEWs)**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Know the Health Extension Worker (HEW) | 42 | 2,239 | 38 | 2,701 | 0.002 |
| Visited by a HEW at home in the past 3 months | 31 | 948 | 27 | 1,025 | 0.017 |
| During the last home visit, the HEW talked about breastfeeding, child feeding or nutrition | 65 | 298 | 63 | 272 | 0.575 |
| Had contact with a HEW in the community, outside their home and outside the health post | 28 | 948 | 23 | 1,025 | 0.003 |
| Compared to pre-COVID times frequency of home visits by Hew have increased | 64 | 298 | 62 | 272 | 0.624 |
| During the last time they met the HEW in the community, HEW talked about breastfeeding, child feeding or nutrition | 19 | 2,238 | 16 | 2,697 | 0.004 |
| Visited the health post in the last 3 months | 31 | 2,238 | 29 | 2,697 | 0.136 |
| During the last visit to the health post, received any advice or information on breastfeeding, child feeding or nutrition | 70 | 370 | 68 | 376 | 0.521 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise.

***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

About 7% of the PSNP households are member of health development army (HDA) and about 16% know a HDA member. Among the PSNP households that know a HDA member, 38% report having had contact with them in the last 3 months and 34% reported having had home visits. When they met the HDA a large majority reported that they talked about breastfeeding, child feeding and nutrition. Non-PSNP households are less likely to be a HDA member (5% vs 7%, $p< 0.01$) and to know a HDA member (14% vs 16%, $p<0.05$) as compared to PSNP households.

**Table 8.6: Interaction with Health Development Army**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Member of a Health Development Army (HDA) | 7 | 2,238 | 5 | 2,697 | 0.002 |
| Know a HDA/leader working in your area? | 16 | 2,238 | 14 | 2,697 | 0.032 |
| Had contact with the HDA/leader in the past 3 months | 38 | 354 | 40 | 368 | 0.567 |
| Been visited at home by a HDA/leader ever | 34 | 354 | 36 | 368 | 0.576 |
| Visited by a HDA/leader at home in the past 3 months | 75 | 119 | 75 | 131 | 0.997 |
| During the last time the HDA/leader visited at home, they spoke about breastfeeding, child feeding or nutrition | 83 | 89 | 79 | 98 | 0.431 |
| Had contact with a HDA/leader in the community, outside your home and outside the health post, in last 3 months | 33 | 354 | 27 | 368 | 0.085 |
| During the last time they met the HDA/leader in the community, they spoke about breastfeeding, child feeding or nutrition | 72 | 115 | 75 | 98 | 0.705 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

## 8.6 Child labor

In this subsection we present summary statistics on work performed by children, grouped into 7-14 years old and 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 8.7. We observe that about 21% PSNP households reported any children being engaged in agricultural activities in the last 7 days as opposed to 26% reporting this among non-PSNP households, the difference being statistically significant. There is also a statistically significant difference in the average hours spent on agricultural activities in the last 7 days among PSNP and non-PSNP households (4.5 hours vs 5.4 hours, $p<0.05$) . A very small percentage of households, ranging from 1-3%, report children being involved in other work activities. Additionally, PSNP households are more likely to report their children being involved in PWs as compared to non-PSNP households (4% vs 0.7%, $p<0.01$), however the proportion of households that report this is small.

**Table 8.7: Child labor among 7-14 years old in the 7 days prior to survey - end-line (% of Households)**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean[1] | N | Mean[1] | N | |
| Engaged in agricultural activities | 21 | 1,560 | 26 | 1,872 | 0.004 |
| Total hours spent in ag activities in last 7 days[2] | 4.5 | 1,560 | 5.4 | 1,872 | 0.039 |
| Engaged in non-ag activities | 3 | 1,560 | 3 | 1,872 | 0.571 |
| Total hours in non-ag activities in last 7 day[2] | 0.68 | 1,560 | 0.68 | 1,872 | 0.997 |
| Engaged in casual labor | 2 | 1,560 | 2 | 1,872 | 0.775 |
| Total hours in casual labor in last 7 day[2] | 0.23 | 1,560 | 0.29 | 1,872 | 0.546 |
| Engaged in salaried work | 0.7 | 1,560 | 0.9 | 1,872 | 0.511 |
| Total hours in salaried work in last 7 day[2] | 0.07 | 1,560 | 0.12 | 1,872 | 0.256 |
| Engaged in PW | 4 | 1,560 | 0.7 | 1,872 | 0.000 |
| Total hours in PW in last 7 days[2] | 0.31 | 1,560 | 0.09 | 1,872 | 0.001 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise. [2] Number.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

Next, we present the descriptive statistics for 15-17 years old children in Table 8.8. We observe that about 27% PSNP households reported any children being engaged in agricultural activities in the last 7 days as opposed to 26% reporting this among non-PSNP households, though the difference is not statistically significant. The only statistically significant difference between PSNP and non-PSNP households is observed whether children are engaged in salaried work in the last 7 days. Also, 5.2% of PSNP households report some of their 15-17 year-olds were engaged in PW, respectively.

**Table 8.8: Child labor among 15-17 years old in the 7 days prior to survey - end-line (% of Households)**

| | PSNP | | Non-PSNP | | P-value of difference by PNSP status |
|---|---|---|---|---|---|
| | Mean | N | Mean | N | |
| Engaged in agricultural activities | 27 | 854 | 26 | 983 | 0.910 |
| Total hours spent in ag activities in last 7 days[2] | 4.77 | 854 | 4.61 | 983 | 0.751 |
| Engaged in non-ag activities | 4.6 | 854 | 4.1 | 983 | 0.600 |
| Total hours in non-ag activities in last 7 day[2] | 0.92 | 854 | 0.89 | 983 | 0.909 |
| Engaged in casual labor | 2.1 | 854 | 1.3 | 983 | 0.193 |
| Total hours in casual labor in last 7 day[2] | 0.17 | 854 | 0.22 | 983 | 0.676 |
| Engaged in salaried work | 1.9 | 854 | 0.5 | 983 | 0.006 |
| Total hours in salaried work in last 7 day[2] | 0.28 | 854 | 0.13 | 983 | 0.162 |
| Engaged in PW | 5.2 | 854 | 0.4 | 983 | 0.000 |
| Total hours in PW in last 7 days[2] | 0.49 | 854 | 0.05 | 983 | 0.000 |

***Notes***: P-values are from a t-test for of equality of means of the outcome across the sub-samples.[1] Percent 'yes' unless specified otherwise. [2] Number.
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

## 8.7 Impact on mid-upper arm circumference among children and women

In this section, we present the results on mid-upper arm circumference (MUAC) among children under 5 years old and women under 40 years old. The mean MUAC among children is 13 and the mean MUAC among women is 23. About 17% and 14% children under 5 years among PSNP households and non-PSNP households, respectively, are categorized as moderately acute malnourished (MAM) (MUAC between 11-12.5cm), difference is statistically significant. About 7.4% and 7.6% children under 5 years are categorized as severe acute malnourished (SAM) (MUAC less than 11cm) among PSNP and non-PSNP households, respectively. There is no statistically significant difference between PSNP and non-PSNP households except for the prevalence of MAM (Table 8.9). The impact estimates do not show an impact of the PSNP on MUAC among children or women. We also do not observe an impact on the prevalence of MAM or SAM among children.[15]

**Table 8.9: MUAC among children and mothers by the Treatment and Control samples in 2021**

| | PSNP households | | | Non-PSNP households | | | P-value | Normalized difference |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | | |
| Child average MUAC | 13.35 | 2.27 | 1,263 | 13.41 | 2.14 | 1,619 | 0.461 | -0.028 |
| % Children suffer from Moderate Acute Malnutrition | 16.6 | 37.2 | 1,690 | 13.7 | 34.4 | 2,130 | 0.014 | 0.122 |
| % Children suffer from Severe Acute Malnutrition | 7.4 | 26.2 | 1,690 | 7.6 | 26.4 | 2,130 | 0.850 | 0.010 |
| Mother average MUAC | 23.39 | 3.02 | 1,482 | 23.51 | 2.89 | 1,875 | 0.274 | -0.038 |

***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 data from the PSNP IV End-line (2021) Household Survey.

**Table 8.10: Impact results from NNM - Child and Mothers' MUAC**

| | Treatment effect (SE) from NNM | N | Comparison mean |
|---|---|---|---|
| Child average MUAC | 0.018<br>(0.086) | 2,875 | 13.408 |
| Child suffers from Moderate Acute Malnutrition | 0.026<br>(0.016) | 2,875 | 0.181 |
| Child suffers from Severe Acute Malnutrition | 0.003<br>(0.012) | 2,875 | 0.100 |
| Mother average MUAC | -0.064<br>(0.109) | 3,356 | 23.506 |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$
***Source***: Authors' computation using data from the PSNP IV End-line (2021) Household Survey.

[15] It should be noted that the impact evaluation was not designed to measure/detect impacts on child or women's nutrition outcomes. The sample was not selected to be able to assess these impacts.

## 8.8 Summary

- This chapter provides descriptive results to understand **RQ4: Has PSNP improved human development outcomes?**
- 30% of the women were ever pregnant since May 2016 and about 10% were pregnant at the time of end-line survey.
- 43% of the women were working on PW when they learned about the pregnancy. About 60% of these women also reported that someone from their household (other than themselves) was working on PW when they learned about the pregnancy.
- 63% of the women who were working of the PW when they found out about their pregnancy they stopped working. Among the women who did not stop working 18% continued to work until they gave birth.
- On average, women stopped working when they were 4 months pregnant and about 45% wished they could have stopped earlier.
- About 30% reported receiving direct support benefits after the women stopped working. And about 24% reported that other household members worked more to compensate for the woman not working.
- 64% report returning to PW after giving birth. On average women went back after 7 months of giving birth and 68% of these women said that going back to work affected their ability to breastfeed their child.
- Among those that were pregnant in the last 3 years, 28% were visited by a health worker/community worker and about 50% reported receiving antenatal care during that pregnancy.
- About 31% reported being pregnant during the COVID period and about 21% reported receiving antenatal care during the pregnancy. We do not observe significant differences between PSNP and non-PSNP households.
- Women's knowledge of breastfeeding is better than timely introduction of complementary foods. The only difference between PSNP and non-PSNP mothers is in breastfeeding knowledge. 65% of PSNP mothers know about immediate initiation of breastfeeding as compared to 62% of non-PSNP mothers. And 61% of PSNP mothers know about exclusive breastfeeding till 6 months as compared to 58% of non-PSNP mothers. Although these differences are significant, the absolute differences are not large.
- 57% of the children under 5 years had a health card (about 56% among PSNP households and 58% among non-PSNP households). However, a much smaller percentage of children were measured About 17%, 16% and 20% were measured for weight, height and mid-upper arm circumference among PSNP households, respectively. It appears that children among PSNP

households are slightly more likely to be measured as compared to those that belong to non-PSNP households.

- On average, 40% of the respondent reported knowing the health extension worker (HEW) – PSNP households more likely than non-PSNP households to know the HEW (42%vs 38%,$p<0.01$). Among those who know the HEW, about 30% were visited by the HEW at home in thepast 3 months – PSNP households are more likely to report this than non-PSNP households (31%vs 26%, $p<0.05$).
- During the last home visit, 63% reported that the HEW talked about breastfeeding, child feeding and nutrition. Compared to pre-COVID times, 64% of the PSNP respondents reported that the frequency of HEW home visit increased.
- About 7% of the PSNP households are member of health development army (HDA) and about 16% know a HDA member.
- About a quarter of households reported that some child members (7-17 year olds) did engage in agricultural activities in the last 7 days. Both the fraction of PSNP households reporting such participation and the average length of participation they report for children in the 7-14 years age group are lower compared to non-PSNP households. The difference, though not large, is statistically significant.
- Some PSNP households report that their children in both the 7-14 and 15-17 age groups did participate in PWs. Nevertheless, both the fraction of households (4% for 7-14 year olds and 5.2% for 15-17 year olds) and the average length of participation (respectively 0.31 hours and 0.49 hours in the last 7 days) are small.
- The mean MUAC among children is 13 and the mean MUAC among women is 23. The impact estimates do not show an impact of the PSNP on MUAC among children or women.

# 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.[16] 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

[16] 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.

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**

| 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] | **0.3 SD=0.58 months reduction in food gap** | **0.4** SD=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 versus 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. 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 PSNP-3 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 phase-4 evaluation; a panel of households is followed in the Highlands over three survey rounds (2016, 2018, 2020) and three repeated cross-sectional surveys (2016, 2018, 2020) 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[17]. 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**

| | |
|---|---|
| **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.

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

# Appendix B: Attrition Analysis

As noted in Chapter 2, the attrition rate for end-line households in Amhara, Oromia and SNNP was 7.3 percent. Using household characteristics measured at the baseline, we estimated a probit model of the correlates of attrition from baseline to midline. Results are shown in Table B1 below.

We find that female headed households and younger households are more likely to attrit from the sample. Somewhat surprisingly, households with more land are significantly more likely to attrit. Households with higher quality housing are less likely to attrit than others. Also, households that experienced a non-drought shock in the last two years are more likely to attrit, but those living in a kebele that received humanitarian assistance in response to a drought in 2015 are less likely to have been dropped from the sample, presumably because they are less likely to have moved. Finally, households in Oromia and SNNP are more likely to attrit than those in Amhara.

**Table B1: Correlates of attrition from baseline to end-line**

| Variables | Attrited |
|---|---|
| Female headed household | 0.213*** |
| | (0.071) |
| Household head's age | -0.034*** |
| | (0.008) |
| Household head's age squared | 0.000*** |
| | (0.000) |
| Number of males | -0.013 |
| | (0.025) |
| Number of females | -0.033 |
| | (0.024) |
| Number of members 16 to 60 | -0.135*** |
| | (0.038) |
| Total land holdings in hectare | 0.052** |
| | (0.023) |
| HH productive asset PCA | -0.024* |
| | (0.013) |
| Housing in moderate to very good condition | -0.236*** |
| | (0.062) |
| Household has corrugated metal roof | -0.145** |
| | (0.066) |
| Household experienced any non-drought shocks in the last 2 years | 0.135** |
| | (0.056) |
| Community is connected to a road made of stone | 0.027 |
| | (0.066) |
| Community is connected to a dirt road | 0.184*** |
| | (0.069) |
| Road is accessible in rainy season | 0.183*** |
| | (0.065) |

| | |
|---|---|
| Number of Development Agents in the kebele | -0.030 |
| | (0.022) |
| Kebele received temporary drought/humanitarian relief in 2015 | -0.226*** |
| | (0.062) |
| Region: Oromia | 0.585*** |
| | (0.089) |
| Region: SNNP | 0.671*** |
| | (0.081) |
| Constant | -0.620*** |
| | (0.221) |
| Number of observations | 5,461 |

***Notes***: * $p<0.10$; ** $p<0.05$; *** $p<0.01$

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

# Appendix C: Heterogeneous Impact Results

In this appendix, we present heterogenous treatment effects. We would like to examine whether those households that were worse off at baseline (in terms of food gap and livestock holdings) experienced a differential impact of participating in the PSNP as compared to those who were better off. To undertake this analysis, we divide the sample using baseline median values of the food gap (and TLU) and estimate the impact of PSNP in each of these subsamples on the outcomes of interest.

## C.1 Heterogenous impacts by baseline levels of food gap

Reviewing the results shown in Table Ax.1, we observe some heterogenous impacts on consumption expenditures and livestock holdings. Households that had a food gap greater than the median food gap at baseline, i.e., households that were worse off at baseline, experienced no impact of the PSNP on consumption expenditure. In contrast, those households that had a food gap lower than the median food gap at baseline experienced a negative impact on consumption expenditures. It also appears that the positive impact on TLU holdings is concentrated among households whose food gap at baseline was lower than the median food gap.

**Table C1: Heterogenous impacts by baseline levels of food gap**

| | FGAP > median FGAP at baseline | | | FGAP <= median FGAP at baseline | | |
|---|---|---|---|---|---|---|
| | Treatment effect (SE) from NNM | N | Comparison mean | Treatment effect (SE) from NNM | N | Comparison mean |
| Change in food gap | 0.158 | 2,177 | -1.849 | -0.021 | 2,739 | 1.387 |
| | (0.152) | | | (0.105) | | |
| Change in Household Dietary Diversity Score (HDDS) | 0.078 | 2,180 | 0.356 | -0.038 | 2,739 | 0.351 |
| | (0.105) | | | (0.103) | | |
| Change in per capita real total consumption expenditures | -10.191 | 2,390 | 87.177 | -50.237** | 2,958 | 101.845 |
| | (24.330) | | | (23.139) | | |
| Change in per capita real food expenditures | -2.879 | 2,282 | 86.968 | -43.332* | 2,872 | 101.127 |
| | (23.588) | | | (22.329) | | |
| Change in per capita real total nonfood expenditures | -9.032* | 2,390 | -5.638 | -2.794 | 2,958 | -5.770 |
| | (4.638) | | | (4.938) | | |
| Change in total tropical livestock units | 0.151 | 2,179 | 0.361 | 0.334*** | 2,739 | 0.060 |
| | (0.162) | | | (0.126) | | |
| Forced to sell any productive assets for food needs | -0.000 | 2,180 | 0.055 | -0.008 | 2,739 | 0.047 |
| | (0.011) | | | (0.009) | | |
| Forced to sell any productive assets for emergency cash needs | 0.013 | 2,180 | 0.050 | 0.003 | 2,739 | 0.042 |
| | (0.010) | | | (0.009) | | |
| Change in income diversification, baseline to end-line | 0.027 | 2,180 | -0.132 | 0.024 | 2,739 | -0.062 |
| | (0.030) | | | (0.026) | | |
| Change in subjective poverty | -0.036 | 2,173 | 0.506 | -0.000 | 2,735 | 0.190 |
| | (0.056) | | | (0.057) | | |
| Total hours worked, 7 days | 4.741*** | 2,180 | 24.558 | 7.040*** | 2,739 | 26.905 |
| | (1.409) | | | (1.452) | | |
| No. activities HH engaged, 7 days | 0.402*** | 2,180 | 0.875 | 0.380*** | 2,739 | 0.887 |
| | (0.040) | | | (0.037) | | |
| Percent hours HH engaged in ag activities last 7 days | -0.209*** | 1,515 | 0.788 | -0.163*** | 1,935 | 0.799 |
| | (0.020) | | | (0.018) | | |
| Percent hours HH engaged in non-ag activities last 7 days | -0.042*** | 1,515 | 0.112 | -0.009 | 1,935 | 0.112 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.013) | | | (0.012) | | |
| Percent hours HH engaged in casual labor last 7 days | -0.000 | 1,515 | 0.059 | -0.016** | 1,935 | 0.048 |
| | (0.010) | | | (0.008) | | |
| Percent hours HH engaged in salaried work last 7 days | 0.007 | 1,515 | 0.030 | -0.006 | 1,935 | 0.034 |
| | (0.007) | | | (0.007) | | |
| Percent hours HH engaged in PW last 7 days | 0.245*** | 1,515 | 0.011 | 0.194*** | 1,935 | 0.008 |
| | (0.013) | | | (0.012) | | |
| HH engaged in ag activities in last 7 days | -0.020 | 2,180 | 0.617 | 0.016 | 2,739 | 0.634 |
| | (0.022) | | | (0.020) | | |
| HH engaged in non-ag activities in last 7 days | -0.005 | 2,180 | 0.125 | 0.026* | 2,739 | 0.130 |
| | (0.015) | | | (0.014) | | |
| HH engaged in casual labor in last 7 days | 0.016 | 2,180 | 0.080 | -0.006 | 2,739 | 0.072 |
| | (0.013) | | | (0.011) | | |
| HH engaged in salaried work in last 7 days | 0.023** | 2,180 | 0.039 | 0.018** | 2,739 | 0.039 |
| | (0.009) | | | (0.009) | | |
| HH engaged in PW in last 7 days | 0.388*** | 2,180 | 0.013 | 0.325*** | 2,739 | 0.012 |
| | (0.016) | | | (0.016) | | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

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

### C.2 Heterogenous impacts by baseline levels of TLU

Reviewing the results shown in Table Ax.2, we observe some heterogenous impacts on food gap, consumption expenditures and income diversification. Households that had TLU holdings less than than the median TLU holdings at baseline, i.e., households that were worse off at baseline, experienced decline in food gap as a result of their participation in the PSNP. Whereas those households that had TLU holdings higher than the median TLU holdings at baseline experienced no impact on food gap. Those with low levels TLU holdings experienced no impact on consumption expenditure while households with TLU holdings greater than the baseline median had a negative impact on consumption expenditure. We also observe a positive impact on income diversification among households that had TLU holdings lower than the baseline median TLU holdings.

**Table C2: Heterogenous impacts by baseline levels of TLU**

| | TLU <= median TLU at baseline | | | TLU > median TLU at baseline | | |
|---|---|---|---|---|---|---|
| | Treatment effect (SE) from NNM | N | Comparison mean | Treatment effect (SE) from NNM | N | Comparison mean |
| Change in food gap | -0.343** | 2566 | 0.207 | 0.090 | 2,355 | -0.006 |
| | (0.143) | | | (0.146) | | |
| Change in Household Dietary Diversity Score (HDDS) | 0.168* | 2568 | 0.304 | -0.203* | 2,356 | 0.395 |
| | (0.101) | | | (0.109) | | |
| Change in per capita real total consumption expenditures | -24.515 | 2,705 | 84.890 | -63.623*** | 2,513 | 101.214 |
| | (26.797) | | | (20.830) | | |
| Change in per capita real total food expenditures | -18.504 | 2,580 | 78.659 | -52.918*** | 2,451 | 103.853 |
| | (25.737) | | | (20.544) | | |
| Change in per capita real total nonfood expenditures | -4.348 | 2,705 | -2.584 | -7.692* | 2,513 | -6.825 |
| | (5.413) | | | (4.182) | | |
| Change in total tropical livestock units | -0.050 | 2,567 | 1.029 | 0.138 | 2,356 | -0.508 |
| | (0.095) | | | (0.182) | | |
| Forced to sell any productive assets for food needs | -0.010 | 2,568 | 0.067 | 0.002 | 2,356 | 0.036 |
| | (0.010) | | | (0.009) | | |
| Forced to sell any productive assets for emergency cash needs | 0.006 | 2,568 | 0.057 | 0.010 | 2,356 | 0.036 |
| | (0.010) | | | (0.010) | | |
| Change in income diversification, baseline to end-line | 0.062** | 2,568 | -0.142 | -0.021 | 2,356 | -0.046 |
| | (0.029) | | | (0.028) | | |
| Change in subjective poverty | -0.006 | 2,562 | 0.387 | -0.046 | 2,351 | 0.262 |
| | (0.055) | | | (0.060) | | |
| Total hours worked, 7 days | 6.180*** | 2,568 | 22.864 | 7.317*** | 2,356 | 28.451 |
| | (1.232) | | | (1.714) | | |
| No. activities HH engaged, 7 days | 0.401*** | 2,568 | 0.878 | 0.416*** | 2,356 | 0.885 |
| | (0.037) | | | (0.041) | | |
| Percent hours HH engaged in ag activities last 7 days | -0.180*** | 1,772 | 0.755 | -0.185*** | 1,682 | 0.825 |
| | (0.019) | | | (0.019) | | |
| Percent hours HH engaged in non-ag activities last 7 days | -0.034*** | 1,772 | 0.114 | -0.017 | 1,682 | 0.110 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.012) | | | (0.012) | | |
| Percent hours HH engaged in casual labor last 7 days | -0.016 | 1,772 | 0.074 | -0.004 | 1,682 | 0.036 |
| | (0.011) | | | (0.007) | | |
| Percent hours HH engaged in salaried work last 7 days | -0.006 | 1,772 | 0.047 | 0.003 | 1,682 | 0.020 |
| | (0.008) | | | (0.006) | | |
| Percent hours HH engaged in PW last 7 days | 0.237*** | 1,772 | 0.010 | 0.202*** | 1,682 | 0.008 |
| | (0.013) | | | (0.014) | | |
| HH engaged in ag activities in last 7 days | 0.008 | 2,568 | 0.598 | -0.002 | 2,356 | 0.651 |
| | (0.020) | | | (0.022) | | |
| HH engaged in non-ag activities in last 7 days | 0.005 | 2,568 | 0.127 | 0.020 | 2,356 | 0.129 |
| | (0.014) | | | (0.016) | | |
| HH engaged in casual labor in last 7 days | 0.015 | 2,568 | 0.089 | 0.007 | 2,356 | 0.064 |
| | (0.012) | | | (0.011) | | |
| HH engaged in salaried work in last 7 days | 0.014 | 2,568 | 0.053 | 0.029*** | 2,356 | 0.028 |
| | (0.010) | | | (0.009) | | |
| HH engaged in PW in last 7 days | 0.360*** | 2,568 | 0.012 | 0.362*** | 2,356 | 0.014 |
| | (0.015) | | | (0.018) | | |

***Notes***: Estimates from the common support sample selected by the matching model. * $p<0.10$; ** $p<0.05$; *** $p<0.01$

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

# Appendix D: 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.

*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.