# The Effect of Ethiopia's Productive Safety Net Program on Livestock Holdings of Rural Households

**Tesfaye Abate Zewdu**

Master thesis for the Master of Philosophy in Economics

Program Option: Research

DEPARTMENT OF ECONOMICS

UNIVERSITY OF OSLO

May, 2015

# The Effect of Ethiopia's Productive Safety Net Program on Livestock Holdings of Rural Households

#### Summary

Food insecurity and vulnerability to poverty is a chronic issue in Ethiopia as the majority of the country's population depends on agriculture for their livelihood. The recurring lack and a high variability of rainfall causes persistent shocks of droughts which forces households to disinvest in assets and leave poor farming families without food crops which can be in turn a cause for the famine of millions of people in the country. To address the severe challenges of food insecurity and poverty, and abolishing recurrent famines in the country, emergency food aid has been taken as a solution for a long period of time. Programs, such as Food for Work and Employment Generation Scheme, were also used as social protection programs in the country since 1980's. However, since the severe drought of 2002/03 brought extreme hunger in the country, the government of Ethiopia, in collaboration with a consortium of donors, has decided to supplement the existing response system with a more predictable and longer-term solution for reducing poverty and vulnerability to food insecurity. Hence, Food Security Program with a component called Productive Safety Net Program [PSNP] was launched in 2005 as a social protection program which makes people's livelihoods more secure. After the inception of the PSNP, there are a number of studies done to evaluate its impact on different outcomes in Ethiopia. However, the novelty of this study can be explained by the facts that the study has used longitudinal national data set collected by Young Lives, which contains pre-intervention information and includes more waves after the start of the program, and little is done on the area of the research. Therefore, this study was done to evaluate the effect of the Public Works (PW) component of Ethiopia's PSNP on livestock holdings of the rural households and investigate its impact disparities across the regions, sex of the household head and drought experience of households.

This study uses Young Lives longitudinal household level data set of Ethiopia collected in three waves, 2002, 2006, and 2009. As far as analysis is concerned, both descriptive and econometric methods were used. Descriptive statistics (mean, percentage, range) was used to summarize the variables in the model. Econometric models, logit model, for the estimation of propensity scores, and matching with difference-in-difference were employed to estimate the effect of the program on livestock holdings and livestock accumulation measured in Tropical Livestock Unit.

In this study, before rushing to interpret the model outputs, effort was made to test the balancing condition of observed covariates between the participant and non-participant groups. The ability of the matching approach to balance the relevant covariates of these two groups was checked by using the standardized bias approach. To avoid very poor matches the common support condition was imposed using the “minima and maxima” comparison approach. The balancing test for various matching methods was conducted for both the full sample and subsamples (by drought, regions, and sex of the household head).

The PSM estimate result shows that participation in the PW component of the PSNP enhanced livestock accumulation; but was not statistically significant. However, the result from the matching with DID estimator reflects that participation in the program had a significant effect on the change in livestock accumulation in TLU. Specifically, the change in livestock holdings of participant households equals 0.57 TLU between 2002 and 2009. The disaggregation results confirmed that the PW payment had a significant effect on the change in livestock accumulation for both subsamples of participant households that were affected by drought and that were not affected by the shock. There was also impact disparity across regions. Results from the matching with DID reflect that the program had a positive effect for all regions but this is statistically significant only for Tigray region. In addition, the there was a substantial impact disparity between female-headed and male-headed participant households. The result from both PSM and matching with DID estimator portrayed that the effect of the program on livestock accumulation was found to be statistically significant for male-headed participant households only. That is, participating male-headed households have benefited from the program in terms of livestock holdings as compared to female-headed participant households. Generally, findings of this study confirmed that participating in the PW component of the PSNP enhances livestock accumulation with considerable effect disparities across regions, sex of the household head, and drought experience of households.

# Preface

First and foremost, I praise the Almighty God, for providing me this opportunity and granting me the capability to proceed successfully. "All things were made by him; and without him was not anything made that was made" (John 1:3).

I would like to express my deepest gratitude to my supervisor, Professor Monique de Haan for her patient guidance, kindness, understanding and continuous support throughout my thesis. I have been extremely lucky to have supervisor who cared so much about my work, and who responded to my frequent questions and queries so promptly.

Besides my advisor, I would like to thank Professor Edwin Leuven for the absolute support to the thesis.

My special thanks goes to the Norwegian State Educational Loan Fund for financing my study through Quota Scheme scholarship. I am also so grateful to the Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India, Peru and Vietnam, for providing me the data used in this study with free of cost.

Finally, I would like to express my gratitude to my parents for their comprehensive support and encouragement throughout my study.

Tesfaye Abate Zewdu

Oslo, Norway

May, 2015

# Table of Contents

## List of Tables

# List of Appendices

# 1. Introduction

Ethiopia is recorded as one of the fastest growing economies in sub-Saharan Africa, with an average gross domestic product [GDP] growth rate of 8.2% between 2000 and 2011(Nganwa, 2013; Deutsche Bank Research, 2013). In addition, using the Ethiopian Household Income and Consumption Expenditure [HICE] survey of the year 2010/11, the Ethiopian Ministry of Finance and Economic Development [MoFED] estimated the proportion of poor people in the country to be 29.6%, falling from 38.7% in 2004/05(MoFED, 2013).

Despite this progress, Ethiopia remains one of the poorest countries in the world in which millions of people are still living in poverty (United Nation Development Program, 2013; MoFED, 2012; World Bank, 2011; UNICEF, 2012). The United States Agency for International Development [USAID] (2014) estimated that 3.76 million people required emergency assistance between the months of August and December in the last decade. The challenge of food insecurity is more severe in the rural areas, where the majority of the population of the country resides. To overcome the problem of food insecurity in the country, emergency food aid has been taken as a solution for a long period of time. The country was one of the largest aid recipients in the world for the past two decades (Little, 2008). However, even if a large proportion of the population has been surviving on imported food aid for many decades, it could not address the underlying causes of food insecurity in the country since food aid has been characterized by low predictability (since it depends on the willingness of donors), poor timing, insufficient assistance for individual beneficiary and highly exposed to corruption (Wiseman *et al.,* 2010, Andersson, Mekonnen, & Stage, 2011; Gilligan, Hoddinott, & Taffesse, 2009; Wiseman *et al.,* 2010); it was often unsuccessful in protecting livelihoods, and it was not cost effective from the donors (such as United States, Canada, Australia, and Japan) point of view (Wiseman, van Domelen, & Coll-Black, 2010). Other programs, such as Food for Work [FFW], which was implemented from 1980 to 2003, and Employment Generation Scheme [EGS], which started in 1997, were used as an instrument to implement food aid as a social protection program in the country (Woldehanna, 2009). The EGS program started as temporary employment scheme for food insecure households and was considered a direct contribution to the rebuilding of household assets, contributing to reduce Ethiopia’s chronic food insecurity (Woldehanna, 2009).

However, since the severe droughts of 2002/03 brought extreme hunger in the country, another program, called Food Security Program [FSP],[1] was announced to supplement the existing response system, the food aid program implemented via EGS, with a more predictable and longer-term solution for reducing poverty and vulnerability to food insecurity (Ethiopian Ministry of Labor and Social Affairs, 2012). In collaboration with a consortium of donors, the GoE launched a component of FSP known as Productive Safety Net Program [PSNP] in 2005 as a social protection program which makes people's livelihoods more secure (Ethiopian Ministry of Agriculture and Rural development, 2006; Woldehanna, 2009). The PSNP is the second largest social protection program in sub Saharan Africa next to South Africa (Gilligan *et al.*, 2009), and operates as a safety net for the long-term by targeting transfers to the needy households for a predictable period of time through its two components, the Public Works [PW] and the Direct Support [DS] (MoADR, 2006; Gilligan *et al.*, 2009). The PW part of the PSNP provides employment opportunities for food insecure households that have able-bodied family member(s) to participate in productive activities, such as rehabilitating land and water resources and developing community infrastructure, including building schools and clinics and rural road rehabilitation. On the other hand, the DS provides unconditional transfer to chronically food insecure households that cannot provide labor to public activities and have no other means of support (MoADR, 2006).[2]

Several empirical studies have been conducted to examine the effect of social protection programs, such as PSNP, on different outcomes. The study done by Gilligan *et al.* (2009) shows that PSNP had a significant effect on consumption. Similarly, Slater, Ashley, Tefera, Buta and Esubalew (2006) also reported that PSNP improved consumption status, asset protection and buildings of participants. A recent study has found that the PW component of the PSNP had a significant effect on households' food security status, improved number of children’s meals consumed and livestock holdings (Berhane, Hoddinott, Kumar, & Tafesse, 2011). Another study conducted by Andersson *et al.* (2011) indicates that the PSNP increased tree holdings but had no effect on livestock holdings. Moreover, the program also has effects on child’s time spent on schooling and work. For instance, the study done by Woldehanna (2009) reveals that the

[1] The FSP consists of the PSNP; Household Asset Building Program [HABP], the Voluntary Resettlement Program [VRP], and the Complementary Community Investment Program [CCI] (MoLSA, 2012).

[2] These components are clearly discussed in section 2.1.1

program increased girls' time spent on studying, it increased child's work for pay, and it decreased child's time spent on child care and household chores.

Even though there are many studies done by different scholars on different outcomes, these studies are based on recall data to construct the baseline data set and were conducted at the early stage of the program. Gilligan *et al.* (2009) used recall data to fill the gap of lack of pre-intervention data. This recall data was collected from the same respondents by employing retrospective questions about demographic characteristics, prior experiences with emergency assistance, assets, and selected food security outcomes such as the size of the food gap. However, respondent recall is often inaccurate since it is hard to remember all past events correctly, resulting in over or under reporting of past events that leads to recall bias (Sudman & Bradburn, 1973). Moreover, the results of the previous findings, such as Andersson *et al.* (2011), may not be the long term impact of the program since transfers were delayed during the first year of implementation of the PSNP (Gilligan *et al.*, 2009), and for the food insecure households, that are in general liquidity constrained, investing on livestock within a short period of time might be a challenging activity. On top of that the external validity of the study of Andersson *et al.* is questionable since the study area, South Wollo, in Amhara region, has been affected by severe droughts repeatedly and, hence, is an impoverished and risky part of the country, even when compared to other low-income areas of rural Africa (Little, Stone, Mogues, Castro, & Negatu, 2006). The current study, therefore, extends the investigations and fills the gap by using a national panel data set from 3 waves, where the first wave is a pre-intervention survey and the last survey was held 4 years after the start of the program.

Thus, the main objective of this study is to investigate the effect of Ethiopia's PSNP on livestock holdings of the rural households. Specifically, this thesis aims to answer the following research questions. 1) Does the PW have a significant effect on livestock capital? 2) Does PW protect livestock holdings in times of shocks (such as drought)? 3) Does the effect of the program vary across the regions and gender of the household head? These research questions are answered by employing the propensity score matching technique and matching with difference-in-differences.

The remaining parts of the thesis are structured as follows. The next section presents a short description of Ethiopia's PSNP including its definition, objectives and eligibility criteria. Section

3 describes the theoretical linkages between consumption smoothing, livestock holdings and the PSNP. Section 4 focuses on the review of previous related empirical studies on PSNP. Section 5 contains a brief description of methodology of the study, which particularly includes the data source and the methods of estimation. Section 6 presents the main empirical results and discussion of the findings. Finally, Section 7 contains the conclusion of the study and its policy implications.

# 2 Description of Ethiopia's Productive Safety Net Program

This section comprises a short description of the PSNP including its definition, objectives and components, targeted individuals, criteria that have been considered in the selection process and transfers mechanisms.

## 2.1 Definition, Objective and components of the Productive Safety Net Program

Before the inception of the PSNP, programs such as emergency food aid, Food-for-Work[FFW], and Employment Generation Scheme[EGS] programs were used as social protection programs in Ethiopia (Woldehanna, 2009). However, after the severe droughts of 2002/03 that brought extreme hunger in the country, the government of Ethiopia has decided to supplement the existing response system, the EGS, with a more predictable and longer-term solution for reducing poverty and vulnerability to food insecurity. To this end, the Food Security Program [FSP] with different components was launched in 2005 (Ethiopian Ministry of Agriculture and Rural development, 2006; 2012; Woldehanna, 2009).[3] In particular, in collaboration with development partners,[4] the government of Ethiopia initiated a component of FSP called the Productive Safety Net Program [PSNP] in the same year as a social protection program which makes people's livelihoods more secure (Wiseman *et al.*, 2010; MoARD, 2006).

The PSNP is designed to the needy households with primary goals as smoothing consumption, protecting asset depletion in times of shocks and thereby encouraging asset accumulation of eligible households that live in chronically food insecure woredas,[5] and creating community assets through public works (MoARD, 2006). Strategically, the program plays a role as an ex-ante alleviation of various shocks, such as drought, by encouraging the rural transformation

[3] The FSP consists of the PSNP; Household Asset Building Program [HABP], the Voluntary Resettlement Program [VRP], and the Complementary Community Investment Program [CCI] (MoLSA, 2012).

[4] According to Wiseman *et al.* (2010), the donors of Ethiopia`s PSNP are: the Canadian International Development Agency, Danish International Development Agency, the European Union, Irish Aid, the Netherlands , Swedish International Development Agency, United Kingdom's Department for International Development, United States Agency for International Development, World Food Program, and World Bank.

[5] A woreda is an administrative district managed by locally elected government. It consists of kebeles, the lowest administrative unit in Ethiopia. Thus, kebeles constitute woreda and woredas in turn constitute zones which in turn build Regional State, called Region.

process (rural diversification), avoiding long-term consequences resulting from the transitory consumption scarcities, enabling households to involve in off-farm activities (invest in a small business productive investment), and promoting market development by scaling up the purchasing power of households (MoARD, 2006).

Furthermore, PSNP plans to bring asset accumulation through its combination of another component of FSP, namely Other Food Security Program [OFSP] which included a suite of activities designed to support agricultural production and food security, and facilitate asset accumulation. Particularly, it included access to credit, assistance in obtaining livestock, small stock or bees, tools, seeds, and assistance with irrigation or water-harvesting schemes, soil conservation, and improvements in pasture land (Berhane et al., 2011). However, since there was a lack of clear eligibility criteria and consequently regional disparities of the OFSP implementation, the government of Ethiopia redesigned and substituted it with Household Asset Building Program [HABP] in 2010 with goals of diversifying income sources and increasing productive assets for its clients, food-insecure households in chronically food insecure woredas (Berhane et al., 2011).

The PW component creates employment opportunities for chronically food insecure households that have able-bodied family member(s) that are required to work five days per month during the agricultural slack season in productive activities, such as rehabilitating land and water resources (such as terracing) and developing community infrastructure, including building schools and clinics and rural road rehabilitation(Sharp, Brown, & Teshome, 2006). In other words, in the PW component, the eligible households are required to participate in the public activities (listed above) in exchange for transfers. On the other hand, the DS component provides unconditional transfers to chronically food insecure and labor constrained households that have no other means of support (Wiseman *et al.,* 2010; MoARD, 2006).

## 2.2 Eligibility of Households and Selection Criteria for the PSNP

The beneficiaries of the program are the food insecure households living in chronically food insecure woredas.[6] The beneficiaries of both the PW and DS components were selected based on both administrative criteria and community knowledge (Brhane *et al.*, 2011; Wiseman *et al.*, 2010). Accordingly, the household is said to be eligible and selected to the PW component, if the household has able-bodied family member(s) within the category of chronically food insecure woredas. The PSNP's Project Implementation Manual [PIM] defines a chronically food insecure household as: "having faced continuous food shortages (usually 3 months of food gap or more) in the last 3 years and received food assistance prior to the commencement of the PSNP; having suddenly become more vulnerable as a result of a severe loss of assets and unable to support themselves for the last 1-2 years; and without family support and other means of social protection and support" (MoRAD, 2006, P.3). Thus, households that meet these criteria are considered as eligible households and will be beneficiary of the program based on the quota given to each woreda and kebele. Similarly, beneficiaries in the DS are also chronically food insecure households that do not have labor to participate in public activities, and do not have reliable and sufficient support from sons/daughters, and relatives. Eligible households for the DS component can also include households whose bread earners are elderly or disabled, pregnant women, lactating mothers and sick individuals (Wiseman *et al.*, 2010).

Generally, in the first program evaluation report of Sharp *et al.* (2006), it is documented that beneficiary households are resource-poor and vulnerable to shocks, unable to produce self-sufficient food even at times of normal rains in the country, and were selected using vulnerability ranking criteria such as household's productive assets, ox and land, and level of poverty. It also indicated that relative poverty was the key selection criteria for both participants and non-participants. Similarly, the recent impact evaluation study done by Berhane *et al.* (2011) also shows that the PSNP has been well targeted to the chronically poor households that engage in activities which generate low returns and are mainly pursued by poor people. Compared to their

[6] A woreda is considered chronically food insecure if it is in one of the 8 regions, namely Tigray, Amhara, Oromia, SNNP, Afar, Somali, rural Harare and Dire Dawa, and has been a recipient of food aid for a significant period, generally for at least each of the last 3 years (MoARD, 2006).

non-participant counterparts, participants, especially beneficiaries of the DS component, have been poorer in both incomes and assets, and cultivated less land.

In order to achieve the above objectives of the program, all the Ministries (particularly the MoFED and MoARD) and their departments, regional Bureaus, woredas and kebeles have specific and integrated roles. At the community level, Community Food Security Task Force [CFSTF][7] is responsible for the identification and registration of the PSNP beneficiaries. The CFSTF identifies the households eligible for public works or direct support transfers, displays the proposed participants' list for a week for public comment and endorsement by the village assembly, and then transfers the finalized list to the Kebele Food Security Task Force [KFSTF] for verification and approval. The KFSTF considers any complaints and takes action where appropriate, and then forwards a compiled list of households to the woreda level for finalization and approval; and once the woreda administration investigates and solves appropriate complains (if any), submission of the final list to the regional Bureau of Agriculture will be conducted (Wiseman *et al.*, 2010; MoARD, 2006).

Initially, the PSNP aimed to cover more than 263[8] woredas in four major regions of the country, namely Tigray, Amhara, Oromia and Southern Nations Nationalities and People's [SNNP], that had been significant recipients of food aid between 2002 and 2004 and operates as a safety net by providing transfers to 4.5 million beneficiaries via either PW or DS **(**Gilligan *et al.*, 2009**).** A recent report shows that around 7.8 million eligible households in the country are enrolled in the program (Nganwa, 2013). The PSNP provides a minimum of five days of payment per month for six months during the agricultural slack season for at least the next five years. A member of the targeted household for the PW employment gets 50 birr[9] (US$2.80) or 15kg of grain per month (Sharp *et al.*, 2006). The type of transfer can be in kind (food), in cash or can be a combination of both; it depends on the transfer that the donors have made. However, the transfers are set at a level intended to smooth out household consumption or fill the food gap over the annual lean

[7] The CFSTF comprised of a kebele official, the local Development Agent (DA) and elected villagers representing men, women, youth, and the elderly (MoARD, 2006)

[8] In 2008, the program operated in 290 of 670 food-insecure woredas (Coll-Black, Gilligan, Hoddinott, Kumar, Taffesse, & Wiseman, 2011).

[9] Birr is the name of Ethiopian currency.

period.[10] However, due to the high inflation rate, adjustments to the wage rates were made over the period of the program and participants received 8 and 10 birr per day in 2008 and 2010, respectively (Brhane *et al.*, 2011).

Since the main objective of PSNP is to safeguard a minimum level of food consumption and enhance livestock accumulation of the needy households, the beneficiaries are expected to "graduate"[11] from the program once they have achieved better livelihoods and become food secure. The support from another component of FSP, namely HABP that provides agricultural extension and credit services in order to diversify income sources and increase productive assets of the participants will also continue after "graduation" (Gilligan *et al.*, 2009; Berhane *et al.*, 2011).

[10] Lean period refers to a period of time in which households forced to food aid; in Ethiopia it mostly occurs between the months of July and September i.e., during a period of planting crops (MoARD, 2006).

[11]The graduation from the PSNP was defined in the 'Graduation Guidance Note' as follows: "A household has graduated when, in the absence of receiving PSNP transfers, it can meet its food needs for all 12 months and is able to withstand modest shocks" (MoARD, 2007, P.2). The guide note also indicates the seven core principles for the introduction and use of benchmarks as well as 16 steps that regions, woredas, kebeles, and communities should undertake in identifying graduates

# 3 Theoretical linkages between Livestock holdings, Consumption Smoothing and the PSNP

In the presence of borrowing constraints and shortage of rainfall, livestock plays a significant role in achieving food security of rural households in developing countries (Deaton, 1991; Rosenzweig and Wolpin, 1993; Webb *et al.*, 1992; Hoddinott, 2006). Deaton (1991) demonstrates that households subject to credit constraints are able to smooth consumption with relatively low asset holdings, i.e., households who do not have access to credit services in times of shocks, their consumption smoothing could be achieved at the cost of livestock holdings, dissaving. Similarly, Zimmermana and Carter (2003) noted that in a resource-poor environment, households save both in the form of conventional buffer assets, such as grain stocks and other safe savings instruments, and in the form of productive assets, like land and livestock, which can be used as a self-insurance consumption smoothing mechanism when income is stochastically variable and credit markets are incomplete. Another study done by Rosenzweig and Wolpin (1993) show that bullocks in India are not only used as source of power in agricultural production, but can also be sold to smooth consumption in time of adverse income shocks. In addition to the short term role of livestock as consumption smoothing device, it also has an important role in agricultural activities as a factor of crops production and provides manure which enhances food security in the long run (Bradford, 1999).

In Ethiopia, livestock also play an important role in achieving food security by providing meat and milk, manure which keeps soil fertility, and generating income (by selling and renting them). Credit constrained rural households also use their livestock as a coping response in times of shocks by selling them (ebremedhin, Hoekstra, & Jemaneh, 2007). However, since livestock are also important factors of production (agricultural activities) in the rural areas, today's livestock depletion in times of adverse transitory shocks leads to loss of crop production efficiency which in turn might be a cause for vulnerability of households to food insecurity in the next period (Webb *et al.*, 1992). Deaton (1991) also demonstrates that households subject to credit constraints are able to smooth consumption with relatively low asset holdings. This indicates that given credit constraints, the exiting livestock asset is depleting as long as households do not get other safety nets for their consumption smoothing.

# 4 Empirical Literature Review

Several empirical studies have been conducted to examine the effect of social protection programs, such as PSNP, on various households' welfare outcomes. Evidence from Alderman and Yemtsov (2012) shows that 62% of the households that participated in the PSNP avoided selling assets in states of food shortages, and 36% avoided using savings to buy food. In addition, they found that 23% of participants acquired new household assets, 46% used healthcare more, and 39% sent more children to school while 50% kept them in school longer.

A study done by Gilligan *et al.* (2009) assessed the impact of the PSNP, on its own and together with the OFSP, on household food insecurity, consumption levels, agricultural and non-farm production enhancement, and asset accumulation by classifying the treatment households in to three categories[12]. The result shows that the definition of participants does matter for the impact of the program. Accordingly, if all participants of PSNP were included in the analysis, the overall effect was insignificant. On the other hand, the beneficiary households that received at least half of the intended transfers experienced a significant improvement in food security. Most importantly, for those households who participated in both the PSNP and OFSP[13], the result indicated a significant effect on food intake and no evidence of deterrence effects in terms of labor supply or private transfers, slower asset growth, than for non-participants[14]. However, Gilligan *et al.* (2009) used recall data to fill the gap of lack of pre-intervention data. This recall data was collected from the same respondents by employing retrospective questions about demographic characteristics, prior experiences with emergency assistance, assets, and selected food security outcomes such as the size of the food gap. However, respondent recall is often inaccurate since it is hard to remember all past events correctly, resulting in over or under reporting of past events that leads to recall bias (Sudman & Bradburn, 1973).

Berhane *et al.* (2011) estimated the impact of Ethiopia's PSNP and other related transfers(OFSP/HABP) on food security using panel data of the Ethiopian Central Statistical

[12] It includes, households who receiving any payment by participating in only PSNP; those households who received at least half of the intended transfer from PSNP (more than 90 Birr per household member), and households who participated in both PSNP and Other Food Security Program [OFSP]

[13] The OFSP/HABP consists of productivity enhancing transfers or services, namely credit, agricultural extension services, technology transfer (including advice on food crop and livestock production, cash cropping, and soil and water conservation), and irrigation and water harvesting schemes(Gilligan *et al.*, 2009; MoARD, 2006).

[14] This negative program impact could be an indicator of selection bias due to unobserved variables or existence of anticipation (Dehejia & Wahba, 1999).

Agency (CSA), called the Ethiopian Food Security Surveys, collected in 2006, 2008 and 2010 from woredas across the four major regions of Ethiopia, namely Tigray, Amhara, Oromiya and SNNP. The estimator that Berhane *et al.* employed was the 'dose–response' model. Instead of considering participation in the program as a binary treatment variable, they considered treatment as a continuous variable, years of receipt of the PSNP. They argue that using binary matching techniques to study the impact of the PSNP is less attractive when there has been considerable movement in and out of the program which makes the construction of control group more difficult. Moreover, the levels of participation in the program may also vary widely. However, these problems are not the issue in the current study since participant households in my data set were beneficiaries of the PW payments both in the second and third waves of the survey after the intervention of PSNP. In addition, knowing the participation status of the households before the second wave (2006) is not significant since, due to lag of implementation, the program had no impact after a year of its inauguration in 2005 (Woldehanna, 2009). The results of Berhane *et al.* show that food security of beneficiaries of both PSNP and the OFSP significantly increased. They also found that the joint effect PSNP and OFSP on livestock holdings is statistically significant and larger than the effect of PSNP alone.

Similarly, using the same data and estimation approach used in Berhane *et al.* (2011), Hoddinott *et al.* (2012) evaluated the impact of the Ethiopia's PSNP and other related transfers (OFSP/HABP) on agricultural productivity. The results of Hoddinott *et al.* indicate that access to both the PSNP and OFSP programs led to considerable improvements in the use of fertilizer and enhanced investments in agriculture likely to improve agricultural productivity among households receiving both programs. In addition, households receiving OFSP transfers that also participated in the PSNP for a long period had significantly higher yields than OFSP beneficiaries with low levels of PSNP participation. They also found that high levels of transfers in the PSNP program alone had no effect on agricultural input use or productivity and a limited impact on agricultural investments.

Using panel data collected in three waves from 2002 to 2007, Andersson *et al.* **(**2011**)** investigated the impact of the PSNP on livestock and tree holdings in Ethiopia. To estimate the impact of the program, Andersson *et al.* employed a linear regression model by including key covariates which are expected to affect the change of livestock and tree holdings. The results prevailed that the program improved tree holdings but had no significant effect on livestock accumulation, and protection of them in times of shocks. However, since transfers were delayed during the first year of implementation of the PSNP (Gilligan *et al.*, 2009), the results of Andersson *et al.* (2011) may not be the long term impact of the program on the outcome since for the food insecure households, who are in general liquidity constrained, investing in livestock within short period of time might be a challenging activity. On top of that the external validity of the study of Andersson *et al.* is questionable since the study area, South Wollo, in Amhara region, has been affected by severe droughts repeatedly and, hence, is an impoverished and risky part of the country, even when compared to other low-income areas of rural Africa (Little, Stone, Mogues, Castro, & Negatu, 2006). In addition, Woldehanna (2009) estimated the impact of PSNP on child welfare by using Young Lives child level panel data set and a propensity score matching model. The estimated results show that the PW component of the PSNP increases child work for pay; reduces children's time spent on child care, household chores and total hours spent on all kind of work combined; and increases girls spending on studying .

This study will contribute to the existing literature by analyzing the impact of the PSNP on livestock holdings using a longitudinal data set with a base line survey and more waves after the start of the program.

# 5 Methodology

## 5.1 Data Source

This study uses the longitudinal household level data set of Young Lives [YL] in Ethiopia that covers the period from 2002 to 2009. YL is an international study of childhood poverty, involving 12,000 children in four developing countries, namely Ethiopia, Peru, India, and Vietnam. The sample consists of two cohorts of children, younger and older. The 2,000 'index children' in each country in the younger cohort were aged 6 to 18 months on the first survey in 2002, and were resurveyed again at age 4 to 5 in 2006 and most recently aged 7 to 8 years in 2009.There were also 1,000 children from each of the four countries in the older cohort (between the ages of 7.5 and 8.5 in 2002), who were resurveyed at age 11.5 to 12.5 in 2006 and in the third round at age 14.5 to 15.5 in 2009.[15]

In Ethiopia, the data set of YL in general has been collected using child level questionnaires, household and community level questionnaires with the primary objective of analyzing child poverty of 3000 children over time. The samples of these 3000 children were selected from 20 sentinel sites, 8 from urban and 12 from rural areas. The sentinel sites were selected by considering multi-dimensional policy variables and other factors relevant to the project from five regions of the country, namely Tigray, Amhara, Oromia, Addis Ababa, and SNNPR. Specifically, selected sites should capture poor areas, with food deficiency; diversity across regions and ethnicities in both urban and rural areas; and should be cost effective. [16]

The data set includes information on pre-intervention and post-intervention outcomes for both participants and non-participants of the PSNP, which is important in this study. The baseline data was collected in 2002 prior to the start of the PSNP (2005). The second and the last rounds of the survey were done after the implementation of the program in 2006 and 2009, respectively. In the survey, the same individual households were followed overtime to form a panel data set. The household-level data used in this study includes, among others, information about productive assets (such as livestock and land), socio-economic and demographic characteristics of households, shocks, program participation, and other income sources.

[15] (www.younglives.org.uk; Woldehanna, 2008).
[16] For more information, see Outes-Leon and Sanchez (2008).

However, since PSNP is targeted at the rural households, this study only uses the rural households data collected from the four biggest regions of the country (Oromia, Amhara, SNNP and Tigray) in which the program was operating for the first time. In this set of YL data, there were 20 households enrolled in OFSP (of which 16 households also participated in PW of PSNP) that possibly have experienced improved impact on livestock holdings. However, these households have been excluded from the analysis. Furthermore, households participating in the DS component of the PSNP (were 50 in total) have also been excluded from this analysis.[17] As a result, the entire sample size for this analysis contains a balanced panel of 1770 households observed in the years 2002 and 2009.

The overall attrition, including death, of the data over the eight years period was 4.8%. It is low in absolute terms and when compared with attrition rates for other longitudinal studies in developing countries (Outes-Leon and Dercon, 2008). Beside this low attrition, testing whether the attrition is exogenous is important since the non-random attrition leads to attrition bias[18]. However, it is difficult to know whether those households who dropped out from the sample between 2002 and 2006 were enrolled in the PSNP or not. Hence, I only test whether the participation in the program has an effect on the attrition rate by using households that withdrew from the sample between 2006 and 2009. The result from logit model estimation confirms that being a beneficiary is not highly correlated with the probability of attrition.  This indicates that attrition apparently is not the general problem to obtaining consistent estimates of the variables of interest. The analysis in this paper is based on a sample with households that were present in all waves.

[17] This is because they are significantly poorer than non-beneficiaries (Gilligan *et al.*, 2009) and the PW participants (Devereux, Sabates-Wheeler, Tefera, & Taye, 2006), hence, it is difficult to match them to the control group; the levels of transfers received by these households are very low (Gilligan *et al.*, 2009); and the component includes some beneficiaries for a limited period of time (such as pregnant and lactating mothers) (MoARD, 2006) which is challenging for estimation. Thus, in this study, both PW and PSNP are used interchangeably.
[18] See Little and Rubin (1987).

## 5.2 Variable Description

The outcome variable in this study is livestock holdings measured in the Tropical Livestock Unit [TLU]. TLU is a common unit to describe livestock numbers of various species as a single figure that expresses the total amount of livestock present, irrespective of the specific composition. The standard used for 1 TLU is one cattle with a body weight of 250 kg (Andersson *et al.*, 2011). In this study livestock are classified into four groups: draught animals- which includes camels, horses, donkeys, young bulls and buffalos with the weight (TLU) of 1; milk animals which includes cows, heifers, and calves with the weight (TLU) of 0.7; small ruminant animals consists of sheep, goat and pig with the weight (TLU) of 0.15; and other animals such as rabbit and poultry with the weight (TLU) of 0.05 (see Woldehanna, Mekonnen, & Alemu, 2008). Note that livestock expressed in TLU, does not include the number of oxen that the households had. This is because of the assumption that the market value of oxen is relatively expensive as compared to other sorts of animals which are common in the study areas. This makes harder to expect investment on such expensive animals by a liquidity constrained households within such short period of time since the program has been started. Thus, TLU here includes breeding animals such as cattle, back animals (donkey and horse), small animals such as sheep, goats and pigs, and others such as chicken and bee which are relatively liquid and there is a readily market in the vicinity of the households. However, the number of oxen in 2002 is taken as one covariate for the estimation of propensity score.

In this study, the inclusion of covariates in the estimation of propensity scores was based on previous studies, economic theories and institutional knowledge. Thus, observed covariates in the pre-intervention period that are expected to affect both the participation in the PW component and the outcome variable can be generally categorized as household demographic characteristics such (household size, age, sex of the household), educational level of household head, productive assets(land and oxen), economic variables (wealth index, non-farm income aid, remittance), shocks(crop failure and death of livestock ), and regional dummies (Amhara, Oromia, SNNP and Tigray).[19]

[19] For definition and types of each variable, see appendix 1

## 5.3 Method of Analysis

Since participation in the PSNP is not random, the simple difference in livestock holdings between treated and non-treated households will not identify the true impact of the program. Thus, to obtain the exact impact of the program, one would ideally look at the difference between the outcomes for PSNP participant households and the outcomes from the same households had they not participated in the program. However, due to the impossibility of observing the same unit under both treatment and control at the same time (Holland, 1986),[20] the outcomes for the households had they not participated in the program can't be observed. More formally, let $T_i$ be the treatment variable, participation in PW of PSNP, a binary variable which equals 1 if household $i$ was enrolled in the PW and 0 otherwise. The Average Treatment Effect on the Treated [ATT] can be calculated as $ATT = E(Y_1 - Y_0|\ T = 1) = E(Y_1|T = 1) - E(Y_0|\ T = 1)$, where $Y_1$ and $Y_0$ are the potential outcomes of interest, livestock holdings measured in TLU,[21] for a household with and without the program respectively. However, as stated above, $E(Y_0|\ T = 1)$ can't be observed. Thus, unless a proxy control group consisting of non-participant households as similar as possible to participant the households is not constructed, using the livestock holdings of untreated household as an estimate of counterfactual will generate a bias.

If the selection is based on variables that are observable to the researcher, the problem of selection bias can be circumvented by controlling for these variables in the propensity score matching method [PSM] (Rosenbaum and Rubin, 1983). The PSM model, however, hinges on two main identifying assumptions. The first assumption is the conditional independence assumption [CIA],[22] which implies, given the observable covariates (X), the potential outcomes do not depend on treatment status ( $X: y_1, y_0 \perp T|X$ ). However, when estimating the ATT, this assumption can be relaxed to $y_0 \perp T|X$ (Rubin, 1977). The second identifying assumption is the common support or overlap assumption. It states that the propensity score [PS], defined by $p(x) = Pr[T = 1|X] = E(T|X)$, should be between 0 and 1 ($0 < Pr[\mathrm{T} = 1|\mathrm{X}] < 1$) both for the treated and untreated. The intuition is that for each treated household, there is another matched untreated household with similar X (Rosenbaum and Rubin, 1983).

[20] Holland called this problem the "fundamental problem of causal inference" (P. 947).
[21] See section 5.2 about TLU.
[22] Using the term of Rosenbaum and Rubin(1983), this assumption is called 'unconfoundedness'

Estimating the PS removes the problem of dimensionality since the information set which justifies matching on X can also justify matching on the $p(x)$ (Rosenbaum and Rubin, 1983). Thus, if $p(x)$ is the propensity score, then: $X\perp T|p(x)$, which asserts that, conditional on the PS, the distribution of covariates should be the same across the treated and the control groups. The participation model ($p(x) = Pr[T = 1|X] = E(T|X)$), was predicted using a logistic probability model.[23] The model was characterized by including relevant variables based on theory and existing empirical studies such Dehejia and Wahba (1999), Woldehanna (2009), Andersson *et al.* (2011), Sharp *et al.* (2006) and Black and Smith (2004).[24]

In implementing PSM, the propensity score was estimated for each household by running the logit model for participation in the PSNP, $T_i$ , on key pre-intervention covariates that are expected to affect both the participation in the program and the outcome variable, livestock holdings in TLU. Then, individuals in the treated group were matched with those in a comparable group on the basis of their PS using various matching algorithms (for robustness purpose) such as the Nearest-Neighbor Matching, Radius Matching, and Kernel matching.

As indicated in Caliendo and Kopeinig (2005), the Nearest Neighbor [NN] matching method is the most straightforward matching estimator with several variants, such as NN without replacement, NN with replacement, and NN with replacement and caliper. The NN matching without replacement indicates that only a single NN untreated household is considered as a match for a treated household, i.e., it is a one-for-one NN matching variant. On the other hand, in the case of NN matching with replacement, an untreated household can be used more than once as a match for treated households.

It is also suggested to use more than one nearest neighbor ('oversampling') which is important to see the influence of the inclusion of more comparison households for the construction of the counterfactual outcome on the estimated effects (Caliendo and Kopeinig, 2005). This k-NN

[23] It predicts the probability of each household receiving the PW of PSNP as a function of pre-program observed household and community characteristics using a sample of program beneficiaries and non-beneficiaries. Formally, $Pr[T = 1|X] = \frac{1}{1+e^{-(Z_i)}}$, where $Z_i = \beta_i X_i$; $X$ is vector of covariates that affects both participation in the program and livestock holdings in the pre-intervention. Since this function is not linear in both dependent and parameters, it can be linearized by taking the log of the odd ratio (see Gujarati, 2004).

[24] These variables include age, sex and schooling of household head; household size and other demographic characteristics; wealth levels before the program; economic shocks such as drought, pests illness or death, livestock loss; indicators of social networks(such as access to market), asset endowment (livestock and land), and geographical variables such as regional dummies.

matching method, where k > 1, takes more than one nearest neighbor for each treated household. In this matching method, the outcome of each treated household is contrasted to a weighted average of the outcome of the k-nearest neighbors; where an equal weight for each of the k neighbors is assumed, i.e., a simple average of the k-NN outcomes is estimated.

Another matching method used in this study is the radius matching method. This matching method is a variant of caliper matching method, but has an advantage to use not only the k-NN within each caliper but all of the control members within the caliper (Caliendo and Kopeinig, 2005). This means that in the caliper matching method, we focus on only k-nearest neighbors of the treated household if the difference in propensity scores of the treated and the control households is less than the predetermined tolerance, the caliper. In the radius matching, however, the number of control households may be more than k-NN as long as the difference in the propensity scores of the treated and the control households is less than the specified caliper. For instance, in the case of NN matching using k nearest neighbors of control households with a certain caliper, we can take k=3 but there may be other control households within the specified caliper which can be included by the radius matching method.

In the kernel matching method, all treated are matched with a weighted average of all controls with weights that are inversely proportional to the distance between the propensity scores of treated and controls. The advantage of using kernel method is that by using all the information in the control set for each treated household the variance of the estimator is reduced. However, since even control households very far from the treated are used this can increase the bias (Caliendo and Kopeinig, 2005). In this study, the weights are calculated using the Gaussian kernel function, since it one of the widest employed kernel functions.[25]

Thus, by following the approach of Caliendo and Kopeinig (2005), after estimating the propensity score and implementing the matching estimator of choice, the PSM estimator for the ATT can then be written as: $ATT = \frac{\sum_{i\in T}[Y_i - \sum_{j\in M(i)} w_{ij} Y_j]}{N_T}$. This means that the observed outcome $Y_i$ of each treated household $i$ is compared with an estimate of $i's$ counterfactual outcome, $\sum_{j\in M(i)} w_{ij} Y_j$. The estimate of this counterfactual is obtained as the average outcome in the

[25] Each matching method has its own pros and cons; a trade-off between bias and variance (see Caliendo and Kopeinig, 2005).

matching set for household $i, M(i)$ that represents the set of control households which are selected to match with $i$. Different methods choose differently the $M(i)$ and the weights to assign to matched controls, $j$ belonging to $M(i)$. After an estimate of the counterfactual has been obtained for each treated household, the difference in the square bracket is calculated, and finally, the average over the treated is taken, dividing by $N_T$, the number of treated households. For the purpose of improving the overlap (common support) assumption (Heckman, Ichimura, & Todd, 1997), treatment observations whose estimated propensity score was greater than the maximum or less than the minimum of the PS of comparison group were dropped. Similarly, comparison group observations were also dropped if their PS were below the minimum or above the maximum of PS of their counterparts. Most importantly, by using Absolute Standardized Bias [ASB] test and t-test the quality of balancing was also checked by testing that both treatment and comparison observations had the same mean distribution of propensity scores and of covariates in the case of both before and after matching. As suggested by Rosenbaum and Rubin (1985), the ASB is a measure of the average imbalance in each covariate X existing between treated and control units. For each covariate X, the ASB can be defined as the difference of sample means in the treated and matched control subsamples as a percentage of the square root of the average of sample variances in both groups. Generally, the ASB before and after matching is given by $ASB = \left| \frac{\bar{X}_T - \bar{X}_C}{\sqrt{0.5(S_T^2 + S_C^2)}} (100) \right|$ where $\bar{X}_T$ and $\bar{X}_C$ indicate the mean in the treatment and control groups, respectively, $S_T^2$ and $S_C^2$ are the variances in the treatment and control groups, respectively. [26]

The longitudinal nature of the data set data was also exploited to combine the matching method and DID estimator which relaxes the assumption of CIA by allowing for time-invariant unobserved characteristics that represent persistent differences between the treated and control group observations (Heckman, Ichimura, Smith and Todd, 1998).That is, in the case of the PSM approach, the strict CIA indicates there are no covariates left, either observable or unobservable, which are expected to affect both participation in the program and outcome variable simultaneously. However, there could be time invariant unobserved characteristics, such as

[26] As a rule of thumb, balancing is considered acceptable for values of the ASB after matching smaller than 5% (caliendo and Kopeinig, 2005). There are also other approaches used to test the quality of matching. For instance, the joint significance and Pseudo-$R^2$ test (Sianesi, 2004) and the stratification test (Dehejia and wahba, 1999).

assertiveness, which could be a cause for unobserved bias in PSM approach. In this case, the matching method and DID estimator would be a best option since these time invariant unobserved covariates would not lead to a biased result since they are eliminated in the first difference of the matched DID approach. Ravallion (2005) also recommended to use both the combination of PSM with DID approach as a method for avoiding selection bias while conducting impact analysis.

In this paper, the ATT in the matching with DID approach is estimated by following some steps. First, the difference in livestock holdings (in TLU) of each household between 2009 and 2002 is generated, i.e., the change in livestock before and after the intervention of the program is calculated. Second, the control households are constructed by using the above mentioned matching methods to find out the counter factual outcome of the treated households. Lastly, taking the change in livestock holdings, found in the first step, as outcome variable, the ATT is estimated by taking the average difference of the change in livestock holdings between the treated and the matched households. More formally, by using the extension of the approach used in Caliendo and Kopeining (2005) for level outcome, the ATT in the matching with DID can be stated as $ATT = \frac{\sum_{i \in T}[\Delta Y_i - \sum_{j \epsilon M(i)} w_{ij} \Delta Y_j]}{N_T}$, where $\Delta Y_i$ and $\sum_{j \epsilon M(i)} w_{ij} \Delta Y_j$ are the outcome estimates of the treated household $i$ and the $i's$ counterfactual outcome between 2009 and 2002, respectively, $w_{ij}$ is weight attached to each control unit $j$ which is being used as a control unit of the treated unit $i$, $M(i)$ represents the set of control households which are selected to match with $i$ , $N_T$ denotes the number of treated households in the common support group.

Finally, to examine the role of the program on protecting livestock capital in times of shocks (drought), I split the total sample in two subsamples on the basis of the occurrence of drought after implementation of the program, i.e., between 2006 and 2009. The first subsample consists of households that were affected by drought and the second contains households that were not affected by drought. Then, following the same procedure of PSM and matching with DID, the role of the program as a buffer stock in times of shocks was analyzed for the two subsamples. In addition, it is also possible to disaggregate the ATT to estimate heterogeneities in impact of the PSNP across the regions and sex of the household head. To this end, I split the full sample in to six subgroups on the basis of sex of the household head and regions (four regions), and the same models were employed to estimate impact of the program using these subgroups.

# 6 Result and Discussion

## 6.1 Descriptive Statistics

The participation of sample households in the PW component of the PSNP by gender of household head in each region is shown in Table 1. The distribution of participant households was varied across regions since the quota allocated for each region depended on the severity of food insecurity in the region in general and in Woreda and kebelle in particular. On average, 35.42 % of the sample households used in this study participated in the PW program. However, across the regions, it ranged from 8.44% (in SNNP) to 76.47% (in Tigray), which accounted 7.5% and 47.69% of the total participant sample households, respectively. Female-headed households accounted for 13.22% of the sample households. From the total participants in the PW component of PSNP, 81.98% were male-headed households and the remaining 18.02% were female-headed participants, which reflect that male-headed households are more likely to be participants in the PW than female-headed households.

Table 1 Distribution of sample by region, sex of household head and participation

| Households Type | **Region** | | | | **Total** |
|---|---|---|---|---|---|
| | Oromia | Amhara | SNNP | Tigray | |
| | Obs. | Obs. | Obs. | Obs. | Obs. |
| Participants | | | | | |
| Male-headed | 136 | 90 | 41 | 247 | 514 |
| Female-headed | 19 | 36 | 6 | 52 | 113 |
| Total | 155 | 126 | 47 | 299 | 627 |
| Non- Participants | | | | | |
| Male-headed | 237 | 246 | 462 | 77 | 1022 |
| Female-headed | 29 | 29 | 48 | 15 | 121 |
| Total | 266 | 275 | 510 | 92 | 1143 |
| Total | 421 | 401 | 557 | 391 | 1770 |

Source: YL data; own computation
Obs. = observations

As Table 2 indicates, the mean of livestock holdings had grown by 18.8% between the first and the third round. Similarly, for both participants and non-participants, except for SNNP, the average livestock index increased in all regions; it ranged from 34.18% (in Oromia) to 0.45% (in SNNP). Not surprisingly, given the targeting of the program, non-participant households had higher average livestock holdings than their PW beneficiary counterparts. In addition to the variation in the average growth rate across regions,[27] it also reveals that the average growth rate in livestock holdings was higher for the PSNP beneficiary households as compared non-participants. However, it is important to keep in mind that this is a descriptive analysis that exploits only the cross-sectional variation in the data and it does not reveal anything about the underlying causal relationship between participation in the PW part of the PSNP and the outcome variable, livestock holdings.[28]

[27] The analysis of variance indicted that the difference was significant at 1% significant level.
[28] The causal relationship is presented in section 6.2.

Table 2 Mean value of livestock holdings (in TLU) by round, region and treatment status of respondents

| Round (R) | **Region** | | | | | | | | | | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Oromia** | | | **Amhara** | | | **SNNP** | | | **Tigray** | | | | | |
| | treated | Non-treated | Total | treated | Non-treated | Total | Treated | Non-treated | Total | Treated | Non-treated | Total | treated | Non-treated | Total |
| 1 | 1.59 | 3.44 | 2.75 | 1.13 | 2.87 | 2.33 | 1.44 | 2.3 | 2.21 | 2.55 | 3.7 | 2.82 | 1.95 | 2.8 | 2.5 |
| 2 | 2.06 | 3.38 | 2.9 | 1.32 | 3.13 | 2.56 | 1.55 | 2.29 | 2.23 | 3.32 | 4.03 | 3.48 | 2.47 | 2.89 | 2.74 |
| 3 | 2.67 | 4.28 | 3.69 | 1.51 | 3.1 | 2.6 | 1.77 | 2.26 | 2.22 | 3.49 | 4.12 | 3.64 | 2.76 | 3.08 | 2.97 |
| Growth (%) between R1 & R3 | 67.92 | 24.42 | 34.18 | 33.63 | 8.01 | 11.59 | 22.92 | -1.74 | 0.45 | 36.86 | 11.35 | 29.1 | 41.54 | 10 | 18.8 |

Source: YL data; own computation

Table 3 presents the mean value of livestock by sex of household head across the rounds. It shows a slight growth in mean value of livestock holdings between female headed and male headed households over the study period under considered. The variation in mean value of livestock for female headed participants and male-headed participants was, however, statistically significant at 1% level.

Table 3 Mean value of livestock holdings (in TLU) by sex of the household head across round

| Round | Household type | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Female-headed | | | | | | Male-headed | | | | | |
| | Participant | | Non-participant | | Total | | Participant | | Non-participant | | Total | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 1 | 1.21 | 1.65 | 1.58 | 1.77 | 1.47 | 2.65 | 2.21 | 1.79 | 3.18 | 2.65 | 2.82 | 2.33 |
| 2 | 1.20 | 1.50 | 1.65 | 1.88 | 1.78 | 2.52 | 2.75 | 2.12 | 3.03 | 2.53 | 2.86 | 2.33 |
| 3 | 1.35 | 1.43 | 1.59 | 2.09 | 2.06 | 2.34 | 3.07 | 2.23 | 3.26 | 2.80 | 3.09 | 2.6 |

Source: YL data; own computation
SD=standard deviation

Table 4 presents the mean value of the livestock holdings by drought experience of households after the start of the program. It shows that the mean value of livestock holdings somewhat increased for both drought experienced and non-experienced households. However, the simple t-test for mean value difference confirms that there was a significant mean difference of livestock holdings between the groups in each round.

Table 4 Mean value of livestock holdings (in TLU) by drought across round

| Round | Household type | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Drought affected | | | | | | Drought non-affected | | | | | |
| | Participant | | Non-participant | | Total | | Participant | | Non-participant | | Total | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| 1 | 2.34 | 1.77 | 3.7 | 2.72 | 3.04 | 2.4 | 1.83 | 1.8 | 2.82 | 2.55 | 2.52 | 2.4 |
| 2 | 2.89 | 2.22 | 3.53 | 2.4 | 3.22 | 2.34 | 2.20 | 2.0 | 2.70 | 2.50 | 2.55 | 2.37 |
| 3 | 3.19 | 2.45 | 3.82 | 2.55 | 3.51 | 2.52 | 2.49 | 2.0 | 2.87 | 2.81 | 2.76 | 2.6 |

Source: YL data; own computation

## 6.2 Econometric Results

This subsection presents results of the impact of the PW component on livestock capital using both propensity score matching and matching with DID estimators. It also contains the disaggregation of the impact of the program by regions, sex of household head, and household's drought experience after the inception of the program. Before proceeding to impact estimation, multicollinearity[29] between variables was checked and there was no indication of multicollinearity problem between the variables that were considered in the estimation of the propensity scores.

### 6.2.1 Results from Propensity Score Matching

The first step in the application of the PSM model is the estimation of propensity scores. Thus, by including all important pre-intervention observable covariates that are expected to affect both participation in the program and the outcome variable simultaneously, the conditional probability of participating in the program for each household was estimated (see Table 5). After the estimation of propensity scores for both participants and non-participants, the common support condition was checked by focusing on the visual analysis of the density distribution of the propensity score in both groups.[30] The distribution of propensity score for non-participants and participants was found to be skewed to the left and the right respectively.[31] In this paper, to avoid very poor matches the common support condition was imposed using the "minima and maxima" comparison approach. Matching methods such as the nearest neighbor without replacement and with replacement matching (including caliper), the kernel matching, and the radius matching were employed to get the control households that have similar characteristics as the treated households. Except the nearest neighbor without replacement matching method, the difference in the balancing results from the rest matching methods is not large (see Table 7). However, among these matching algorithms, the radius matching method with a caliper of 0.05 resulted in the best balancing. Using this matching method, only 3 treated households (0.48% of

[29] The variance inflation factor (VIF), which shows how much the variance of the coefficient estimate is being inflated by multicollinearity, was applied to detect the multicollinearity problem. As a rule of thumb, when the VIF>10, it indicates the existence of multicollinearity(see Gujirati, 2004, 2004). In addition, if there is a perfect multicollinearity between variables, Stata drop the variable automatically.
[30] The common support condition was also checked by using statistical test called Kolmogorov-Smirnov (KS) test (not reported).
[31] For the common support condition, see the figures in Appendix 2A and 2B.

the whole treated households) were out of the common support group. Therefore, by excluding these 3 treated households, the ATT is redefined as the mean treatment effect for those treated falling within the common support. As indicated in Bryson, Dorsett and Purdon (2002), since the proportion of lost individuals is fairly small, the estimated effect on the remaining households can be viewed as representative.

Table 5 Estimation of propensity scores using a logit model

| Variable | Coeff. | Std.Err. | P>\|z\| |
|---|---|---|---|
| Sex of the household head | 0.099 | 0.435 | 0.821 |
| Age of the household head | -0.001 | 0.007 | 0.876 |
| Wealth index | -3.017 | 0.840 | 0.000 |
| Number of oxen | -0.679 | 0.187 | 0.000 |
| Land size (own) | -0.378 | 0.150 | 0.012 |
| Male family members between age of 18 and 65 | -0.507 | 0.349 | 0.146 |
| Female family members between age of 18 and 65 | -0.049 | 0.491 | 0.920 |
| Dummy for illiterate education level of the head | 2.306 | 0.520 | 0.000 |
| Dummy for elementary education level of the head | 2.393 | 0.523 | 0.000 |
| Dummy for secondary education level of the head | 1.320 | 0.599 | 0.028 |
| Dummy for informal education level of the head | 2.072 | 0.533 | 0.000 |
| Dummy for Amhara region | -1.684 | 0.579 | 0.004 |
| Dummy for Oromia region | -2.797 | 0.647 | 0.000 |
| Dummy for SNNP region | -3.751 | 0.694 | 0.000 |
| Dummy for non-farm income | 0.571 | 0.242 | 0.018 |
| Dummy for aid | 0.244 | 0.140 | 0.083 |
| Dummy for remittance from the family members | 0.286 | 0.150 | 0.056 |
| Dummy for the death of livestock | -0.237 | 0.150 | 0.114 |
| Dummy for crop failure | 0.500 | 0.293 | 0.088 |
| Family members below age of 17 and above 65 | 0.096 | 0.045 | 0.032 |
| Square of own land | 0.065 | 0.023 | 0.004 |
| Square of males between age of 18 and 65 | 0.094 | 0.101 | 0.355 |
| Square of females between age of 18 and 65 | -0.001 | 0.126 | 0.994 |
| Interaction between Amhara and number of oxen | -0.858 | 0.283 | 0.002 |
| Interaction between Amhara and sex of the head | -0.179 | 0.545 | 0.743 |
| Interaction between Oromia and number of oxen | -0.337 | 0.228 | 0.140 |
| Interaction between Oromia and sex of the head | 0.806 | 0.612 | 0.188 |
| Interaction between SNNP and number of oxen | -0.421 | 0.416 | 0.311 |
| Interaction between SNNP and sex of the head | -0.659 | 0.669 | 0.325 |
| Interaction between crop failure and Amhara | 0.115 | 0.397 | 0.772 |
| Interaction between crop failure and Oromia | 0.943 | 0.430 | 0.028 |
| Interaction between crop failure and SNNP | 0.078 | 0.442 | 0.859 |
| Constant | 0.148 | 0.795 | 0.852 |
| Summary statistics | Number of obs = 1770<br>Wald chi2(32) = 469.49<br>Prob > chi2 = 0.0000<br>Pseudo R2 = 0.3463 | | |

Source: YL data: own computation

Another important testable condition in the application of the PSM model is the balancing condition. The key idea of PSM is to construct a control group from the group of non-participant individuals. Most importantly, the balancing condition should also be satisfied; the control group should be as similar as possible to the treatment group with respect to pre-intervention observable characteristics. Thus, in this paper, the ability of the matching approach to balance the relevant covariates of these two groups was checked by using standardize bias approach[32] of Rousbaum and Rubin (1985). The imbalance between the participant and non-participant households before and after matching was estimated[33]. Accordingly, as indicated in Table 6, the overall average bias before and after matching was found to be 19.1% and 2.6%, respectively. This indicates that a significant reduction in bias could be achieved by the matching procedure which in turn implies the improved balancing characteristic in the treatment and the matched comparison group. The matching for each covariate was also very strong; none of the differences in mean of covariates is statistically significant.

[32] The standardized bias shows the percentage reduction in bias among pre-intervention characteristics of the treated and the control groups the, i.e., it reflects how much of the bias was eliminated by matching (Rousbaum & Rubin, 1985).

[33] Matching is implemented with the Stata module PSMATCH2 by Leuven and Sianesi (2003).

Table 6 Balancing test on differences between treated and control households in mean of observed variables before and after matching for the full sample by using radius matching

| Variable | Sample | Mean | | Percentage bias | Absolute bias reduction after matching (%) | t-test | |
|---|---|---|---|---|---|---|---|
| | Unmatched Matched | Treated | Control | | | t | p>\|t\| |
| headsex | Unmatched | 0.84 | 0.91 | -23.0 | | -4.81 | 0.000 |
| | Matched | 0.84 | 0.84 | -0.6 | 97.5 | -0.09 | 0.928 |
| headage | Unmatched | 39.6 | 38.65 | 8.4 | | 1.70 | 0.089 |
| | Matched | 39.60 | 40.23 | -5.8 | 30.9 | -0.95 | 0.344 |
| wi | Unmatched | 0.13 | 0.12 | 10.0 | | 1.91 | 0.056 |
| | Matched | 0.13 | 0.12 | 3.8 | 61.9 | 0.77 | 0.438 |
| anioxen | Unmatched | 0.76 | 0.98 | -25.4 | | -4.85 | 0.000 |
| | Matched | 0.76 | 0.7 | 2.9 | 88.5 | 0.61 | 0.541 |
| ownland | Unmatched | 0.76 | 0.78 | -2.9 | | -0.58 | 0.561 |
| | Matched | 0.76 | 0.8 | -1.8 | 37.2 | -0.32 | 0.745 |
| male1865 | Unmatched | 0.99 | 1.2 | -26.5 | | -5.30 | 0.000 |
| | Matched | 1.00 | 0.99 | 1.1 | 95.7 | 0.21 | 0.835 |
| female1865 | Unmatched | 1.12 | 1.2 | -7.0 | | -1.39 | 0.164 |
| | Matched | 1.12 | 1.10 | 3.8 | 45.1 | 0.71 | 0.480 |
| illitrate | Unmatched | 0.53 | 0.4 | 26.1 | | 5.26 | 0.000 |
| | Matched | 0.53 | 0.55 | -4.3 | 83.4 | -0.76 | 0.448 |
| elem | Unmatched | 0.28 | 0.35 | -15.3 | | -3.05 | 0.002 |
| | Matched | 0.28 | 0.3 | 1.0 | 93.2 | 0.19 | 0.849 |
| secon | Unmatched | 0.02 | 0.07 | -24.8 | | -4.63 | 0.000 |
| | Matched | 0.02 | 0.02 | 2.4 | 90.4 | 0.65 | 0.513 |
| informal | Unmatched | 0.16 | 0.15 | 4.4 | | 0.89 | 0.373 |
| | Matched | 0.16 | 0.15 | 2.8 | 36.7 | 0.49 | 0.627 |
| Amhara | Unmatched | 0.20 | 0.24 | -9.2 | | -1.83 | 0.068 |
| | Matched | 0.20 | 0.22 | -4.7 | 48.5 | -0.84 | 0.399 |
| Oromia | Unmatched | 0.25 | 0.23 | 5.1 | | 1.02 | 0.307 |
| | Matched | 0.25 | 0.27 | -3.9 | 22.1 | -0.68 | 0.499 |
| SNNP | Unmatched | 0.07 | 0.45 | -95.1 | | -17.7 | 0.000 |
| | Matched | 0.08 | 0.08 | -0.3 | 99.7 | -0.07 | 0.946 |
| dummynai | Unmatched | 0.12 | 0.10 | 9.8 | | 2.01 | 0.044 |
| | Matched | 0.12 | 0.12 | 0.4 | 95.6 | 0.07 | 0.942 |
| daid | Unmatched | 0.43 | 0.34 | 19.9 | | 4.03 | 0.000 |
| | Matched | 0.43 | 0.43 | 0.6 | 97.2 | 0.10 | 0.924 |
| remitdum | Unmatched | 0.32 | 0.20 | 25.7 | | 5.29 | 0.000 |
| | Matched | 0.31 | 0.28 | 6.8 | 73.6 | 1.14 | 0.253 |
| lstockdeath | Unmatched | 0.34 | 0.32 | 3.8 | | 0.76 | 0.447 |
| | Matched | 0.34 | 0.35 | -1.6 | 56.6 | -0.29 | 0.775 |
| cropfail | Unmatched | 0.70 | 0.52 | 36.0 | | 7.16 | 0.000 |
| | Matched | 0.70 | 0.68 | 2.5 | 93.1 | 0.46 | 0.647 |
| depen | Unmatched | 2.73 | 2.80 | -4.2 | | -0.84 | 0.401 |
| | Matched | 2.73 | 2.71 | 1.4 | 67.4 | 0.24 | 0.810 |
| **Summary of balancing test** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.332 | 763.73 | 0.000 | 19.1 | 12.6 | 159.1 | 0.78 |
| Matched | 0.004 | 7.23 | 0.996 | 2.6 | 2.4 | 15.2 | 1.05 |

Source: YL data: own computation

Table 7 presents summary balancing test using various matching methods. It indicates that the imbalance between the treated and non-treated is significantly lowered after matching. However, the balancing test from the radius matching with a caliper of 0.05 resulted in the best balancing.[34]

Table 7 Robustness checks for balancing test

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.332 | 763.73 | 0.000 | 19.1 | 12.6 | 159.1 | 0.78 |
| Matched | 0.162 | 281.12 | 0.000 | 12.5 | 10.7 | 102.5 | 1.48 |
| NN matching method using 5 nearest neighbors with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.332 | 763.73 | 0.000 | 19.1 | 12.6 | 159.1 | 0.78 |
| Matched | 0.006 | 11.14 | 0.943 | 3.4 | 3.0 | 18.9 | 0.88 |
| Kernel matching method with a bandwidth of 0.06 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.332 | 763.73 | 0.000 | 19.1 | 12.6 | 159.1 | 0.78 |
| Matched | 0.006 | 10.25 | 0.963 | 2.9 | 2.4 | 18.1 | 1.12 |
| **Radius matching with a Caliber of 0.05** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.332 | 763.73 | 0.000 | 19.1 | 12.6 | 159.1 | 0.78 |
| **Matched** | **0.004** | **7.23** | **0.996** | **2.6** | **2.4** | **15.2** | **1.05** |

Source: YL data: own computation

[34] For the radius matching method, the balancing test for each covariate is presented in table 6, but for the rest matching methods it is available on request by the author.

Table 8 presents the effect of the PW component of the PSNP on livestock holdings and livestock accumulation from various matching methods. It shows that the differences in the effect of the program among the employed matching methods are not large, except the NN matching without replacement. However, since the radius matching resulted in the best balancing, the interpretation is done only from this matching algorithm. Hence, the result in the 1st row portrays that participation in the PW component of the PSNP slightly enhanced livestock accumulation by 0.21 TLU. Nevertheless, the effect is very low and it was not statistically significant.

Table 8 Effect of PW payment on livestock accumulation (in TLU)

| Out come | Matching method | | | |
|---|---|---|---|---|
| | NN(1) | 5-NN (0.05) | **Radius (0.05)** | Kernel |
| TLU | -0.19<br>(0.13)<br>[-1.35] | 0.29<br>(0.21)<br>[1.39] | **0.21**<br>**(0.20)**<br>**[1.04]** | 0.15<br>(0.19)<br>[0.75] |
| ∆TLU | 0.46***<br>(0.11)<br>[4.12] | 0.56***<br>(0.16)<br>[3.59] | **0.57*****<br>**(0.16)**<br>**[3.5]** | 0.58***<br>(0.16)<br>[ 3.63] |

-The outcome TLU and ∆TLU indicate livestock holdings estimated using the PSM, and matching with DID, respectively
-The first column under the matching method indicates nearest neighbor matching using 1 nearest neighbor the second column implies nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, the third column indicates radius matching with a caliper of 0.05, and the last column implies Kernel matching using Normal density function with a bandwidth of 0.06 (the default).
- Values in parenthesis and bracket are standard errors and t-values, respectively. For the kernel matching method, bootstrap standard errors with replication of 100 is used.
*** Significant at 1 % level of significance, ** significant at 5% level of significance, * statistically significant at 10% level of significance.

Source: YL data: own computation

## 6.2.2 Result from Matching with Difference-in-Difference

The PSM estimator was employed by assuming that all characteristics which are expected to affect both participation in the program and the outcome variable are all observable, i.e., taking the strong assumption of CIA in to consideration. However, the CIA assumption might not hold if there remain important unobserved differences between the treated and untreated groups. As a result, the longitudinal data at hand was also exploited to combine the matching method and DID estimator which relaxes the assumption of CIA by allowing for time-invariant unobserved characteristics between the treated and control group observations (Heckman, Ichimura, Smith, & Todd, 1998; Ravallion, 2005).

Table 8 also presents the impact of PW payments on the change in livestock accumulation (in TLU) using the matching with DID estimator. The matching with DID impact estimate (in the 2$^{th}$ row) shows that the effect of PW on the change in livestock holdings equals 0.57 TLU between 2002 and 2009. This indicates that compared to non-participant households, livestock accumulation of participant households was increased by 0.57 TLU. The effect was also statistically significant at 1% level. The possible reason for the significant result from matching with DID estimator could be that participation in the program was not only affected by the observed pre-intervention covariates but also by other unobserved covariates which are assumed to be time invariant. This implies that the estimate from the PSM approach suffered from omitted variable bias. The result of this study is also significantly different from the previous related empirical study done by Anderson *et al.* (2011). Anderson *et al.* found that participation in the PSNP had no significant effect on livestock capital in the study area, South Wollo, in Amhara region. On the other hand, the result of the current study was akin to that of a study conducted by Berhane *et al.* (2011); they reported that participation in the PW component of the PSNP enhanced livestock accumulation significantly.

### 6.2.3 Effect Disaggregation

The disaggregation of the impact of the program was also conducted to see the impact disparities across sex of household head, regions, and drought experience of households. Hence, I create eight subsamples; four subsamples of regions (Amhara, Tigray, Oromia, and SNNP), two subsamples of male-headed and female-headed households, and the remaining two were from drought affected and non-affected households. Accordingly, using the same econometric approach used in the full sample analysis, the propensity scores estimation and balancing test were also conducted for each subsample. However, when all observed covariates (including higher order and interactions terms) that were used in the estimation of propensity scores of the full sample were included in the estimation of propensity scores of each subsample, the balancing test indicates the existence of imbalance between the treated and non-treated households. As a result, the matching method which was applied and resulted in best balancing in the full sample did not work for the subsamples. Thus, it is important to note that the key observed covariates which are expected to affect both the participation in the program and the outcome variable were the same for both the full sample and subsamples in the estimation of propensity scores. The inclusion of the higher order and interaction terms in the estimation of the propensity scores of subsamples was, however, conducted if and only if adding these terms resulted in an improved balancing condition.

Following the estimation of the propensity scores for each observation using each subsample,[35] the balancing test was conducted for various matching methods. Thus, using the subsamples from drought affected areas, the balancing test was checked using nearest neighbor matching without replacement, nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, radius matching with a caliper of 0.05, and kernel matching methods. Among these matching methods, the nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05 resulted in the best balancing.[36] The overall average bias before and after matching, which measures the imbalance between the covariates of the treated and control households, was found to be 17.6% and 9.5%, respectively. However, using this matching algorithm and specification, the balancing test indicates that bias for covariates such as number of dependent family members, dummy for death of livestock, dummy for Oromia region, dummies for illiterate and informal education level of the household head, and wealth index was above 10% after matching, but the imbalance on these covariates was not statistically significant.

Table 9 presents the estimated effect of the PW on livestock holdings and accumulation using various matching methods listed above. However, since the nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05 resulted in the best balancing, the interpretation of the results for drought affected households subsample is based on this matching method. The result from the PSM approach shows that participation in the PW component deteriorated the livestock accumulation; but not statistically significant. In the same Table, the result from matching with DID estimator ($\Delta$TLU) portrays that the participation in the PW had positive effect on livestock holdings. The effect of the PW on the change in livestock holdings equals 0.73 TLU between 2002 and 2009. This indicates that compared to non-participant households, livestock

[35] The result of propensity scores estimation using each subsample is available on request by the author.
[36] The summary result for the balancing test is presented in appendix 3.1. However, the balancing test for each covariate is available on request by the author.

accumulation of participant households that were affected by drought was increased by 0.73 TLU. The effect was also statistically significant at 5% level. However, this result has to be treated with caution since there has been imbalance between the treated and control groups on some covariates.

Similarly, using the same matching methods used in for the drought affected households, the kernel matching resulted in the best matching quality for the subsample of households that were not affected by drought. The overall average bias was 17% and 3.2% before and after matching, respectively. This matching was also very strong; none of the differences was statistically significant.[37]

Table 9 also presents the estimated effect of the PW on livestock holdings and accumulation for households that were not affected by drought. The effect of the program resulted from the PSM approach, depicts that participation in the PW has a small positive effect on livestock holdings. However, the effect was not statistically significant. On the other hand, the result from the matching with DID estimator for the same households indicates that for those participant households that were not affected by drought in 2006 and 2009, the change in livestock accumulation equals 0.42 TLU over the period under consideration. That is, compared to non-participant households, livestock accumulation of participant households that were not affected by drought increased by 0.42 TLU. This effect was also statistically significant at 5% level.

[37] See appendix 3.2 for the summary result of balancing test. However, the balancing test for each covariate is available on request by the author.

Table 9 Disaggregation of effect of the PW on livestock holdings and accumulation by drought

| Sample | Out come | NN(1) | **5-NN(0.05)** | Radius (0.05) | **Kernel** |
|---|---|---|---|---|---|
| Drought affected households | TLU | -0.45<br>(0.23)<br>[-1.94] | **-0.36**<br>**(0.43)**<br>**[-0.83]** | -0.45<br>(0.44)<br>[-1.01] | -0.45<br>(0.43)<br>[-1.04] |
| | ΔTLU | 0.59<br>(0.19)<br>[3.1] | **0.73****<br>**(0.36)**<br>**[2.03]** | 0.77<br>(0.45)<br>[1.72] | 0.76<br>(0.43)<br>[1.77] |
| Drought non-affected households | TLU | -0.26<br>(0.18)<br>[-1.50] | 0.08<br>(0.28)<br>[0.27] | 0.20<br>(0.23)<br>[0.85] | **0.14**<br>**(0.22)**<br>**[0.63]** |
| | ΔTLU | 0.28<br>(0.14)<br>[1.98] | 0.37<br>(0.23)<br>[1.63] | 0.42<br>(0.18)<br>[2.38] | **0.42****<br>**(0.17)**<br>**[2.43]** |

-The outcome TLU and ΔTLU indicate livestock holdings estimated using the PSM, and matching with DID, respectively
-The first column under the matching method indicates nearest neighbor matching using 1 nearest neighbor the second column implies nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, the third column indicates radius matching with a caliper of 0.05, and the last column implies Kernel matching using Normal density function with a bandwidth of 0.06 (the default).
- Values in parenthesis and bracket are standard errors and t-values, respectively. For the kernel matching method, bootstrap standard errors with replication of 100 is used.
*** Significant at 1 % level of significance, ** significant at 5% level of significance, * statistically significant at 10% level of significance

Source: YL data: own computation

Considering the sex of household head, the estimation of propensity scores and the balancing test were also conducted on subsamples of female-headed and male-headed households using the same matching algorithms stated above. For both female-headed and male-headed subsamples, the kernel matching algorithm resulted in the best balancing; and hence the interpretation of the effect of the program for these subsamples hinges on this matching method.[38]

The disaggregation on the effect of the program by sex of the household head is presented in Table 10. The results from both PSM and matching with DID show that participation in the PW payment deteriorated livestock holdings and accumulation for female-headed households. However, this effect was not statistically significant. On the contrary, for male-headed participant households, the result from PSM indicates that participation in the PW increased livestock holdings by 0.25 TLU. However, the effect was not statistically significant. On the other hand, the result from the matching with DID estimator reflects that the change in livestock

[38] Summary results of balancing tests are presented in appendix 4.1 and 4.2 for female-headed and male-headed subsamples, respectively. The results of balancing test for each covariate using both subsamples are available on request by the author.

accumulation for participant male-headed households equals 0.78 TLU between 2009 and 2002. This means that compared to non-participant male-headed households, livestock accumulation of participant male-headed households increased by 0.78 TLU. This result was also statistically significant at 1% level.

Table 10 Disaggregation of effect of the PW on livestock holdings and accumulation by sex of the household head

| Sample | Out come | NN(1) | 5-NN(0.05) | Radius (0.05) | **Kernel** |
|---|---|---|---|---|---|
| Female-headed | TLU | -0.43<br>(0.26)<br>[-1.68] | 0.14<br>(0.34)<br>[0.39] | -0.10<br>(0.38)<br>[-0.27] | **-0.13**<br>**(0.36)**<br>**[-0.38]** |
| | ∆TLU | -0.05<br>(0.23)<br>[-0.11] | 0.06<br>(0.33)<br>[0.18] | -0.15<br>(0.334)<br>[-0.45] | **-0.17**<br>**(0.32)**<br>**[-0.51]** |
| Male-headed | TLU | -0.12<br>(0.15)<br>[-0.80] | 0.47<br>(0.25)<br>[1.91] | 0.34<br>(0.24)<br>[1.40] | **0.25**<br>**(0.23)**<br>**[1.11]** |
| | ∆TLU | 0.60<br>(0.126)<br>[4.79] | 0.80<br>(0.20)<br>[4.1] | 0.76<br>(0.20)<br>[3.76] | **0.78*****<br>**(0.19)**<br>**[4.07]** |

-The outcome TLU and ∆TLU indicate livestock holdings estimated using the PSM, and matching with DID, respectively
-The first column under the matching method indicates nearest neighbor matching using 1 nearest neighbor the second column implies nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, the third column indicates radius matching with a caliper of 0.05, and the last column implies Kernel matching using Normal density function with a bandwidth of 0.06 (the default).
- Values in parenthesis and bracket are standard errors and t-values, respectively. For the kernel matching method, bootstrap standard errors with replication of 100 is used.

*** Significant at 1 % level of significance, ** significant at 5% level of significance, * statistically significant at 10% level of significance.

Source: YL data: own computation

Similarly, using the same matching methods discussed above, the estimation of propensity scores and checking the balancing condition were also conducted. Accordingly, the radius matching method with a caliper of 0.05 resulted in the best balancing for subsamples from Amhara, SNNP and Tigray regions. For Oromia region, among other matching methods, the kernel matching algorithm resulted in the best balancing.[39] Thus, the interpretation of the estimated effect of the program is based on the matching method which resulted in the best balancing for the corresponding subsamples of regions.

Table 11 presents the estimated effect of the PW on livestock holdings and livestock accumulation using subsamples from four regions, namely Amhara, Oromia, SNNP, and Tigray. The results from the PSM approach indicate that participating in the PW deteriorated the livestock holdings in all regions, except Tigray region; but all these results were not statistically significant. However, considering the results from matching with DID estimator, participation in the PW affected livestock accumulation positively for all regions. However, the effect was statistically significant only for Tigray region. That is participating in the PW component resulted in a significant change in livestock accumulation for Tigray region. Specifically, for those participant households from Tigray region, the change in livestock accumulation equals 0.88 TLU, meaning that compared to non-participant households from Tigray region, livestock accumulation increased by 0.88 TLU for participant households. The result was also statistically significant at 1% level. The causes of impact disparities possibly due to implementation efficiency; PSNP is well implemented in Tigray region (Berhane *et al.*, 2011). The significant effect of the program for Tigray region is different from the findings of Berhane *et al.* (2011). They found that the program had no impact for Tigray and Oromia regions, but for Amhara and SNNP the program had a significant effect.

[39] For the summary results of balancing test of subsample from each region, see appendix 5.1, 5.2, 5.3, and 5.4 for Amhara, Oromia, SNNP and Tigray region, respectively. The results of balancing test for each covariate using each region subsample are available on request by the author.

Table 11 Disaggregation of effect of the PW on livestock holdings and accumulation by region

| Sample | Out come | NN(1) | 5-NN(0.05) | **Radius (0.05)** | **Kernel** |
|---|---|---|---|---|---|
| Amhara | TLU | -0.50<br>(0.21)<br>[-2.37] | -0.10<br>(0.32)<br>[-0.26] | **-0.13**<br>**(0.29)**<br>**[-0.44]** | -0.23<br>(0.28)<br>[-0.81] |
| | ΔTLU | 0.30<br>(0.20)<br>[1.48] | 0.12<br>(0.27)<br>[0.45] | **0.25**<br>**(0.24)**<br>**[1.04]** | 0.24<br>(0.23)<br>[1.05] |
| Oromia | TLU | -1.16<br>(0.33)<br>[-3.48] | 0.49<br>(0.64)<br>[0.77] | 0.064<br>(0.45)<br>[0.14] | **-0.05**<br>**(0.42)**<br>**[-0.12]** |
| | ΔTLU | -0.024<br>(0.26)<br>[-0.09] | 0.35<br>(0.5)<br>[0.70] | 0.46<br>(0.40)<br>[1.14] | **0.45**<br>**(0.37)**<br>**[1.21]** |
| SNNP | TLU | -0.55<br>(0.34)<br>[-1.64] | -0.52<br>(0.35)<br>[-1.47] | **-0.362**<br>**(0.232)**<br>**[-1.55]** | -0.39<br>(0.22)<br>[-1.79] |
| | ΔTLU | -0.11<br>(0.35)<br>[-0.33] | -0.08<br>(0.37)<br>[-0.21] | **0.01**<br>**(0.298)**<br>**[0.03]** | 0.08<br>(0.29)<br>[0.29] |
| Tigray [40] | TLU | | 0.20<br>(0.41)<br>[0.48] | **0.26**<br>**(0.39)**<br>**[0.66]** | 0.18<br>(0.40)<br>[0.46] |
| | ΔTLU | | 0.76<br>(0.29)<br>[2.66] | **0.88*****<br>**(0.28)**<br>**[3.18]** | 0.84<br>(0.27)<br>[3.07] |

-The outcome TLU and ΔTLU indicate livestock holdings estimated using the PSM, and matching with DID, respectively

-The first column under the matching method indicates nearest neighbor matching using 1 nearest neighbor the second column implies nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, the third column indicates radius matching with a caliper of 0.05, and the last column implies Kernel matching using Normal density function with a bandwidth of 0.06 (the default).

- Values in parenthesis and bracket are standard errors and t-values, respectively. For the kernel matching method, bootstrap standard errors with replication of 100 is used.

*** Significant at 1 % level of significance, ** significant at 5% level of significance, * statistically significant at 10% level of significance.

Source: YL data: own computation

[40] Note that in Tigray region the number of control households is less than the number of treated households, 93 and 299 respectively. Hence, the nearest neighbor without replacement is impossible.

# 7 Conclusions

The study was set out to investigate the effect of Ethiopia's PSNP on livestock holdings of the rural households. Specifically, the study sought to answer the following research questions: 1) Does the PW have a significant effect on livestock holdings? 2) Does PW protect livestock holdings in times of shocks (such as drought)? 3) Does the effect of the program vary across the regions and gender of the household head?

Propensity score matching techniques and matching with difference-in-differences were used to analyzes the effect of the PW payment on livestock holding in Ethiopia. Various matching methods such as nearest neighbor matching without replacement, nearest neighbor matching using 5 nearest neighbors with a caliper of 0.05, radius matching with a caliper of 0.05, and the kernel matching methods were used for the purpose of robustness check. Albeit the result from the PSM was not statistically significant, it indicates that participation in the PW component of the program had positive effect on the livestock holdings. However, taking the advantage of combining the matching method with difference-in-differences estimator, I found that the program encouraged livestock accumulation among PW component of PSNP beneficiary households as compared to non-beneficiary households. Moreover, I found evidence of large and significant heterogeneities in outcomes depending on experience of shocks (drought), sex of the household head and region of the households. Specifically, taking households' experience of drought shocks, there was evidence that the program had more effect on livestock accumulation for participant households that were affected by drought than participant households that were not affected by the shock after the start of the program. The effect of the program also varies from null effect to a significantly positive effect on livestock holdings and accumulation across regions. Similarly, participating male-headed households have benefited more from the program in terms of livestock accumulation than female-headed participant households. In general, participant households seem to have invested significantly more on livestock capital than non-participants after the inception of the PSNP.

In this study, it is also important to note that one possible limitation of this paper could be the inability to control bias due to externalities to the comparison group, the violation of the 'stable unit treatment value assumption' [SUTVA]. [41] Even when the comparison households are not directly provided with the PW employment, they may indirectly be affected by spillovers from the treatment households. For instance, the public activities done by the treated households lead to community asset development which benefits both the treated and the comparison group. In addition, if households that were shortlisted but not assigned to PSNP in the first phase disinvested livestock capital in expectation of being enrolled in the second phase, one might see some improvements in livestock holdings among these control group households. However, since these rural households are usually highly credit constrained, anticipation effects are expected to be very low.

[41] For detail discussion of SUTVA, also called the "assumption of no general equilibrium effects", see Wooldridge (2010).

# References

Alderman, H., & Yemtsov, R. (2012). Productive Role of Safety Nets: Background Paper for the World Bank 2012-2022, Social Protection and Labor Strategy (Discussion Papers, No.1203). Retrieved from the World Bank website: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2012/03/22/000333037_20120322011452/Rendered/PDF/676090NWP012030Box367885B00PUBLIC0.pdf

Andersson, C., Mekonnen, A., & Stage, J. (2011). Impacts of the Productive Safety Net Program in Ethiopia on Livestock and Tree Holdings of Rural Households. *Journal of Development Economics*, *94*(1), 119–126. doi:10.1016/j.jdeveco.2009.12.002

Berhane, G., Hoddinott, J., Kumar, N., & Tafesse, A. S. (2011). The Impact of Ethiopia's Productive Safety Nets and Household Asset Building Program: 2006-2010. International Food Policy Research Institute. Retrieved from https://www.ids.ac.uk/files/dmfile/Impact_Report_FINAL_JAN2012.pdf

Black, D., & Smith, J. (2004). How Robust is the Evidence on the Effects of the CollegeQuality? Evidence from Matching. *Journal of Econometrics, 121*(1-2), 99-124. doi: 10.1016/j.jeconom.2003.10.006

Bradford, G.E. (1999). Contributions of animal agriculture to meeting global human food demand. *Livestock Production Science 59*(2-3), 95–112. Retrieved from http://dx.doi.org/10.1016/S0301-6226(99)00019-6

Bryson, A., Dorsett, R., & Purdon, S. (2002). The Use of Propensity Score Matching in the Evaluation of Labour Market Policies (Working Paper No. 4). Retrieved from http://eprints.lse.ac.uk/4993/1/The_use_of_propensity_score_matching_in_the_evaluation_of_active_labour_market_policies.pdf

Caliendo, M. & Kopeinig, S. (2005). Some Practical Guidance for the Implementation of Propensity Score Matching (IZA Discussion Paper, No. 1588). Retrieved from http://ftp.iza.org/dp1588.pdf

Coll-Black, S., Gilligan, D. O., Hoddinott, J., Kumar, N., Taffesse, A. S., & Wiseman, W. (2011). Targeting Food Security Interventions When "Everyone is Poor": The Case of

Ethiopia's Productive Safety Net Programme (Ethiopia Strategy Support Program II, Working paper No. 24). Retrieved from http://www.ifpri.org/sites/default/files/publications/esspwp24.pdf

Deaton, A. (1991). Saving and Liquidity Constraints. *Econometrica, 59*(5), 1221-1248.Retrieved fromhttps://www.princeton.edu/~deaton/downloads/Saving_and_Liquidity_Constraints.pdf

Dehejia, R., & Wahba, S. (1999). Causal Effects in Non-expermental Studies: Reevaluating the evaluation of training programs. *Journal of American Statistical Association*, *94*(448),1053 1062. doi: 10.1080/01621459.1999.10473858

Deutsche Bank Research (2013). Sub-Saharan Africa: A bright spot in spite of key challenges. Retrieved from http://www.dbresearch.com/PROD/DBR_INTERNET_EN PROD/PROD0000000000317226.pdf

Devereux, S., Sabates-Wheeler, R., Tefera, M., & Taye, H. (2006). *Ethiopia's Productive Safety Net Program (PSNP): Trends in PSNP Transfer within Targeted Households (FinalReport).* Retrieved from Institute of Development Studies website: https://www.ids.ac.uk/files/PSNPEthiopia.pdf

Gebremedhin, B., Hoekstra, D., & Jemaneh, S. (2007). Heading towards Commercialization? The case of Live Animal Marketing in Ethiopia: Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project (Working Paper No. 5). Retrieved from the International Livestock Research Institution (ILRI) websites: https://cgspace.cgiar.org/bitstream/handle/10568/573/LiveAnimal_Marketing.pdf?sequence=1

Gilligan, D. O., Hoddinott, J., & Taffesse, A. S. (2009). The Impact of Ethiopia's Productive Safety Net Programme and its Linkages. *Journal of Development Studies, 45*(10), 1684-1706. doi:10.1080/00220380902935907

Gujarati, N. D. (2004). *Basic Econometrics* (4th. ed). NY: McGraw-Hill.

Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies*, *64*(4), 605-654. doi: 10.2307/2971733

Hoddinott, J. (2006). Shocks and Their Consequences across and Within Households in Rural Zimbabwe. *Journal of Development Studies*, *42*(2), 301-321. doi:10.1080/00220380500405501

Hoddinott, J., Berhane, G., Gilligan, D., Kumar, N., & Taffesse, A. S. (2012). The Impact of Ethiopia's Productive Safety Net Programme and related transfers on agricultural productivity. *Journal of African Economies*, *21*(5), 761-786. doi:10.1093/jae/ejs023

Holland, P. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, *81*(396), 945-960. doi: 10.1080/01621459.1986.10478354

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Software version 4.0.11 22 oct2014. Retrieved from http://ideas.repec.org/c/boc/bocode/s432001.html.

Little, P.D. (2008). Food Aid Dependency in the North Eastern Ethiopia: Myth or Reality? *World Development,* 36 (*5*), 860-874. doi:10.1016/j.worlddev.2007.05.006

Little, R. J., & Rubin, D. B. (1987). Statistical analysis with missing data, Volume 539. Wiley New York.

Little, P. D., Stone, M. P., Mogues, T., Castro, A. P., & Negatu, W. (2006). 'Moving in Place': Drought and Poverty Dynamics in South Wollo, Ethiopia. *Journal of Development Studies*, 42 (*2*), 200–225. doi: 10.1080/00220380500405287

Ministry of Agriculture and Rural Development (2007). Productive Safety Net Programme: Graduation Guidance Note. Addis Ababa: Ministry of Agriculture and Rural Development.

Ministry of Agriculture and Rural Development (2006). Productive Safety Net Program: Program Implementation Manual (Revised). Retrieved from http://www.cangoethiopia.org/assets/docs/Revised_Ethiopia_Safety_Net_PIM_July_2006.pdf

Ministry of Finance and Economic Development (2012). Ethiopia's Progress towards Eradicating Poverty: An Interim Report on Poverty Analysis Study (2010/11).

Ministry of Finance and Economic Development (2013). Development and Poverty in Ethiopia 1995/96-2010/11.

Ministry of Labor and Social Affair (2012). National Social Protection Policy of Ethiopia (finaldraft). Retrieved from http://phe-ethiopia.org/resadmin/uploads/attachment-188 Ethiopia_National_Social_Protection.pdf

Namara, R. E., Makombe, G., Hagos, F., & Awulachew, S. B. (2008). Rural Poverty and Inequality in Ethiopia: Does Access to Small-scale Irrigation Make a Difference? Retrieved from http://publications.iwmi.org/pdf/H044072.pdf

Nganwa, P. (2013). Resources for poverty eradication: A background paper on Ethiopia. Retrieved from Development Initiatives website: http://devinit.org/wpcontent/uploads/2013/08/Ethiopia-Resources-for-poverty-eradication.pdf

Outes-Leon, I., & Dercon, S. (2008). Survey Attrition and Attrition Bias in Young Lives: Young Lives Technical Note No. 5. Retrieved from http://www.younglives.org.uk/files/technical-notes/survey-attrition-and-attrition-bias-in-young-lives/view

Outes-Leon, I. & Sanchez, A. (2008). An Assessment of the Young Lives Sampling Approach in Ethiopia (Young Lives Technical Note No.1). Retrieved from http://www.younglives.org.uk/files/technical-notes/an-assessment-of-the-young-lives-sampling-approach-in-ethiopia

Ravallion, M. (2005). Evaluating Anti-Poverty Programs. *World Bank Policy Research Working Paper 3625*.

Rosenbaum, P.R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *70*(1), 41–55. doi:10.1093/biomet/70.1.41

Rosenzweig, M. R., & Wolpin, K.I. (1993). Credit Market Constraints, Consumption Smoothing, and Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India. *Journal of Political Economy*, *101*(2), 223–244. Retrieved from http://www.jstor.org/sTable/pdfplus/2138818.pdf

Rubin, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics, 2* (1), 1-26. doi: http://dx.doi.org/10.3102/10769986002001001

Save the Children. (2012). Nutrition in the First 1,000 Days: State of the World's Mothers 2012. Retrieved from http://www.savethechildren.org/atf/cf/%7B9def2ebe-10ae-432c-9bd0df91d2eba74a%7D/STATEOFTHEWORLDSMOTHERSREPORT2012.PDF

Sharp, K., Brown, T., & Teshome, A. (2006).Targeting Ethiopia's Productive Safety Net Program (PSNP) *(Report)*. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odiassets/publications-opinion-files/3966.pdf

Sianesi, B. (2004): An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s. *The Review of Economics and Statistics*, *86*(1), 133-155. doi:10.1162/003465304323023723

Slater, R., Ashley, S., Tefera, M., Buta, M., & Esubalew, D. (2006). PSNP Policy, Program and Institutional Linkages (Report). Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/3964.pdf

Sudman, S., & Bradburn, N. M. (1973). Effects of Time and Memory Factors on Response in Surveys. *Journal of the American Statistical Society*, *68(344)*, 805-815. doi: 10.1080/01621459.1973.10481428

UNICEF. (2012). Situation of Boys and Girls in Ethiopia: Past, Present and Future. Retrieved from http://www.unicef.org/ethiopia/ET_sitan_2012.pdf

United Nation Development Program (2013). The Rise of the South: Human Progress in a Diverse World (Human Development Report 2013). Retrieved from http://hdr.undp.org/sites/default/files/reports/14/hdr2013_en_complete.pdf

Webb, P., Joachim, V. B., & Yisehac, Y. (1992). Famine in Ethiopia: Policy Implications of Coping Failure at National and Household levels (Report No. 92). International Food Policy Research Institute. doi: 0-89629-095-6

Wiseman,W., van Domelen, J., & Coll-Black, S. (2010). Designing and Implementing a Rural Safety Net in a Low Income Setting: Lessons Learned from Ethiopia's Productive Safety Net Program 2005-2009. Retrieved from World Bank website:

http://siteresources.worldbank.org/SAFETYNETSANDTRANSFERS/Resources/Ethiopia PSNPL essonsLearnedLite.pdf

Woldehanna, T. (2009). Productive Safety Net Program and Children's Time use Between Work and Schooling in Ethiopia (Working Paper No. 40). Retrieved from Young Lives website: http://r4d.dfid.gov.uk/PDF/Outputs/YoungLives/YL-WorkingPaper-40.pdf

Woldehanna, T., Gudisa, R., Tafere, Y., & Pankhurst, A. (2011). Understanding Changes in the Lives of Poor Children: Initial findings from Ethiopia (Round 3 Survey Report). Retrieved from http://www.younglives.org.uk/files/country-reports/round-3-country- reports/round-3-survey-report_ethiopia

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data (*2nd ed*).* London: The MIT Press.

World Bank. (2011). Project Performance Assessment Report Ethiopia (No. 62723): Emergency Drought Recovery Project (IDA GRANT-H0280). Retrieved from https://ieg.worldbankgroup.org/Data/reports/PPAR_Ethiopia_Emergency_Drought_Recovery_Project.pdf

Woldehanna, T., Mekonnen, A., & Alemu, T. (2008).*Young Lives: Ethiopia Round 2 Surve Report*. Retrieved from Young Lives website: http://www.younglives.org.uk/files/country-reports/round-2-country-reports/country- report-ethiopia-2008

Zimmermana, F. & Carterb, M. R. (2003). Asset Smoothing, Consumption Smoothing and the Reproduction of Inequality under Risk and Subsistence Constraints. *Journal of Development Economics, 71*(2), 233 – 260. doi:10.1016/S0304-3878(03)00028-2

# Appendices

## Appendix 1 Type, definition and measurement of variables

| Variable | Types and definition | Measurement |
|---|---|---|
| Ti | Dummy, participation in the PSNP | 1 if a household was participant in the PW of PSNP, and 0 otherwise |
| TLU | Continuous , the total livestock capital owned by the household | Tropical Livestock Units |
| headage | Continuous, age of the household head | In years |
| headsex | Dummy, gender of the household head | 1 if male, and 0 otherwise |
| ownland | Continuous, total area of land owned by the household | In hectare |
| female1865 | Continuous, number of female family members between the age of 18 and 65 | Number of female members in the household between the age of 18 and 65 |
| male1865 | Continuous, male family members | Number of male members in the household between the age of 18 and 65 |
| fdepen | Continuous, female dependent | Number of female dependent between the age of 0 and 18 years |
| mdepen | Continuous, male dependent | Number of male dependent between the age of 0 and 18 years |
| depen | Continuous, dependent family members | Number of family members below the age of 18 and above 65 |
| illitrate | Dummy, education level of the household head | 1 if no education, and 0 otherwise |
| informal | Dummy, informal education level of the household head, such as religious education | 1 if she/he has informal level of education and 0 otherwise |
| elem | Dummy, elementary education level of the household head | 1 if she/he has elementary schooling, and 0 otherwise |
| secon | Dummy, secondary education level of the household head | 1 if she/he has secondary schooling, and 0 otherwise |
| postplus | Dummy, post-secondary and above education level of the household head | 1 if she/he has post-secondary and above schooling, and 0 otherwise |
| wi | Continuous, wealth index of the household | The index constructed from three variables (housing quality, consumer durables and service) |
| cropfail | Dummy, crop failure shock | 1 if the household affected by a shock of crop failure for the last three years, and 0 otherwise |
| lstockdeath | Dummy, death of livestock shocks | 1 if the household affected by a shock of livestock loss for the last three years, and 0 otherwise |
| Amhara | Dummy, Amhara region | 1 if the household lives in Amhara region, and 0 otherwise |
| Oromia | Dummy, Oromia region | 1 if the household lives in Oromia region, and 0 otherwise |
| SNNP | Dummy, SNNP region | 1 if the household lives in SNNP region, and 0 otherwise |
| Tigray | Dummy, for Tigray region | 1 if the household lives in Tigray region, and 0 otherwise |
| anioxen | Continuous, oxen owned by the household | In number |

| Appendix 1 cont'd | | |
|---|---|---|
| remitdum | Dummy for remittance | 1 if the household received money from NGOs and individuals, and 0 otherwise |
| dummynai | Dummy for non-farm income | 1 if the household engaged in non-farm activities, and 0 otherwise |
| daid | Dummy for aid | 1 if the household received any aid in the last three years, and 0 otherwise |
| male1865squ | Square of male1865 | Male1865* Male1865 |
| female1865squ | Square of female1865 | female1865* female1865 |
| ownlandsqu | Square of ownland | Ownland*ownland |
| amha_oxen | | Amhara*anioxen |
| Amhara_headsex | | Amhara *headsex |
| oromia_anioxen | | Oromia*anioxen |
| oromia_headsex | | Oromia*headsex |
| Snnp_anioxen | | SNNP*anioxen |
| Snnp_headsex | | SNNP*headsex |
| cropfailAmhara | | Cropfail*Amhara |
| cropfailOromia | | Cropfail*Oromia |
| cropfailSNNP | | Cropfail*SNNP |

Appendix 2A Distribution of propensity score for the full sample

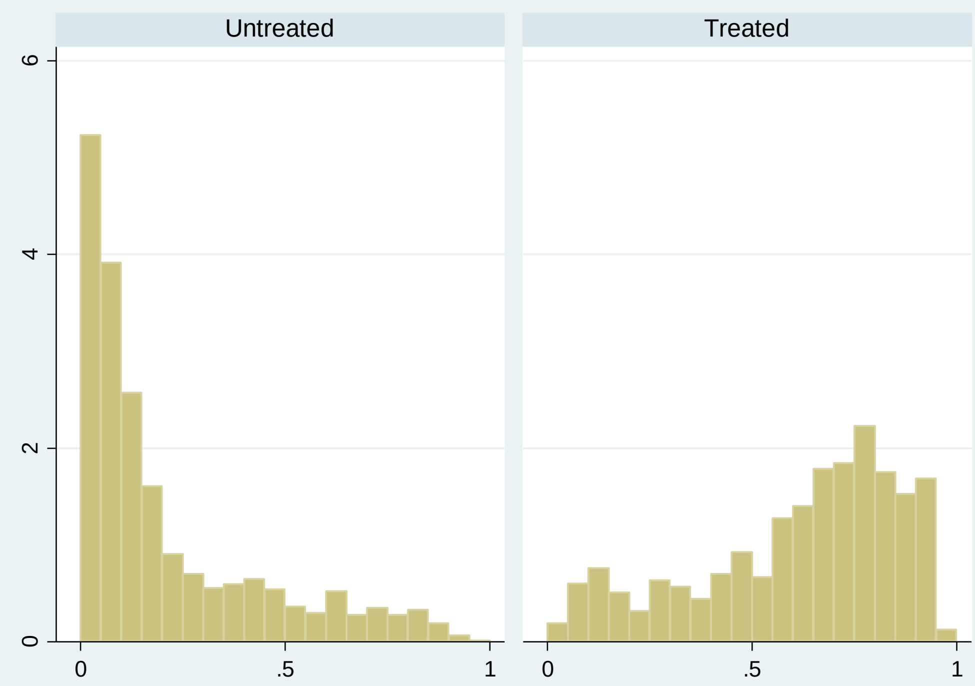


Figure 1.A Distribution of propensity scores of untreated and treated groups

Appendix 2B common support condition for the full sample

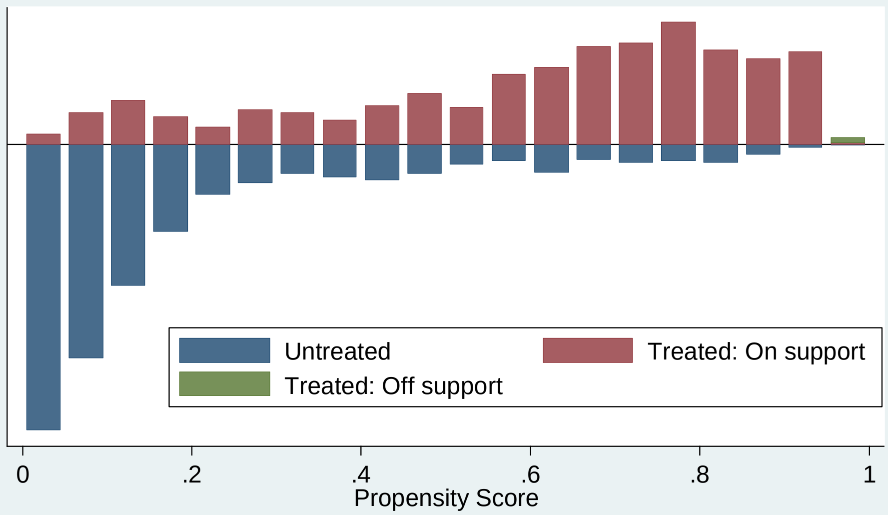


Figure 1B Graphical inspection of common support condition

Appendix 3.1 summary result for the balancing test using various matching algorithms, for drought affected households

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.109 | 75.32 | 0.000 | 17.6 | 14.7 | 80.1 | 1.20 |
| Matched | 0.082 | 52.69 | 0.000 | 15.4 | 11.9 | 69.1 | 1.09 |
| | | | | | | | |
| **5-NN matching with a caliper of 0.05** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.109 | 75.32 | 0.000 | 17.6 | 14.7 | 80.1 | 1.20 |
| Matched | 0.041 | 26.60 | 0.064 | 9.5 | 9.3 | 40.2 | 5.53 |
| Radius matching with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.109 | 75.32 | 0.000 | 17.6 | 14.7 | 80.1 | 1.20 |
| Matched | 0.039 | 25.27 | 0.089 | 9.9 | 10.0 | 46.4 | 2.29 |
| kernel with a bandwidth of 0.06 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.109 | 75.32 | 0.000 | 17.6 | 14.7 | 80.1 | 1.20 |
| Matched | 0.038 | 24.32 | 0.111 | 9.9 | 10.2 | 45.5 | 2.23 |

Source: YL data: own computation

Appendix 3.2 summary statistics for the balancing test using various matching algorithms, for drought affected households

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.127 | 198.29 | 0.000 | 17.0 | 16.1 | 88.9 | 0.67 |
| Matched | 0.031 | 32.49 | 0.013 | 8.1 | 8.8 | 41.9 | 0.91 |
| | | | | | | | |
| 5-NN matching with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.127 | 198.29 | 0.000 | 17.0 | 16.1 | 88.9 | 0.67 |
| Matched | 0.019 | 19.79 | 0.285 | 6.2 | 7.1 | 26.7 | 3.44 |
| Radius matching with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.127 | 198.29 | 0.000 | 17.0 | 16.1 | 88.9 | 0.67 |
| Matched | 0.006 | 6.10 | 0.992 | 3.4 | 2.1 | 18.0 | 0.98 |
| **Kernel with a bandwidth of 0.06** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.127 | 198.29 | 0.000 | 17.0 | 16.1 | 88.9 | 0.67 |
| **Matched** | **0.006** | **6.14** | **0.992** | **3.3** | **1.7** | **18.0** | **1.03** |

Source: YL data: own computation

Appendix 4.1 summary result for the balancing test using various matching algorithms, for female-headed subsample

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.308 | 99.86 | 0.000 | 24.1 | 18.3 | 147.3 | 0.54 |
| Matched | 0.222 | 63.27 | 0.000 | 17.5 | 13.2 | 121.1 | 0.46 |
| NN matching using 5 nearest neighbors with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.308 | 99.86 | 0.000 | 24.1 | 18.3 | 147.3* | 0.54 |
| Matched | 0.101 | 28.33 | 0.077 | 11.5 | 10.4 | 77.3* | 1.42 |
| Radius matching with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.308 | 99.86 | 0.000 | 24.1 | 18.3 | 147.3 | 0.54 |
| Matched | 0.035 | 10.10 | 0.966 | 6.2 | 5.5 | 44.4* | 1.36 |
| **Kernel with bandwidth 0.06** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.308 | 99.86 | 0.000 | 24.1 | 18.3 | 147.3 | 0.54 |
| **Matched** | **0.033** | **9.35** | **0.978** | **6.3** | **4.4** | **42.6** | **1.42** |

Source: YL data: own computation

Appendix 4.2 summary result for the balancing test using various matching algorithms, for male headed subsample

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.352 | 688.89 | 0.000 | 18.2 | 13.0 | 164.5 | 0.82 |
| Matched | 0.167 | 237.54 | 0.000 | 11.5 | 8.5 | 104.2 | 1.75 |
| | | | | | | | |
| NN matching using 5 nearest neighbors with a caliper of 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.352 | 688.89 | 0.000 | 18.2 | 13.0 | 164.5 | 0.82 |
| Matched | 0.014 | 19.22 | 0.508 | 5.0 | 5.6 | 27.6 | 1.13 |
| | | | | | | | |
| **kernel matching with bandwidth 0.06** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.352 | 688.89 | 0.000 | 18.2 | 13.0 | 164.5 | 0.82 |
| Matched | 0.012 | 16.56 | 0.681 | 4.8 | 4.4 | 25.5 | 1.34 |
| | | | | | | | |
| Radius matching with caliper 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.352 | 688.89 | 0.000 | 18.2 | 13.0 | 164.5 | 0.82 |
| Matched | 0.012 | 17.47 | 0.622 | 4.8 | 5.6 | 26.3 | 1.16 |

Source: YL data: own computation

Appendix 5.1 summary result for the balancing test using various matching algorithms, for Amhara region

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.214 | 106.36 | 0.000 | 28.0 | 26.1 | 121.4 | 0.46 |
| Matched | 0.029 | 10.02 | 0.903 | 10.1 | 9.5 | 40.3 | 1.37 |
| | | | | | | | |
| NN matching with replacement and caliper 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.214 | 106.36 | 0.000 | 28.0 | 26.1 | 121.4 | 0.46 |
| Matched | 0.060 | 20.87 | 0.232 | 12.5 | 8.0 | 58.8 | 0.77 |
| **Radius with caliper 0.05** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.214 | 106.36 | 0.000 | 28.0 | 26.1 | 121.4 | 0.46 |
| **Matched** | **0.008** | **2.64** | **1.000** | **3.2** | **2.0** | **20.5** | **0.73** |
| Kernel with bandwidth 0.06 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.214 | 106.36 | 0.000 | 28.0 | 26.1 | 121.4 | 0.46 |
| Matched | 0.009 | 3.20 | 1.000 | 3.7 | 2.5 | 22.5 | 1.08 |

Source: YL data: own computation

Appendix 5.2 summary result for the balancing test using various matching algorithms, for Oromia region

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.246 | 134.69 | 0.000 | 19.8 | 14.7 | 105.2 | 0.12 |
| Matched | 0.074 | 31.14 | 0.013 | 7.7 | 5.3 | 66.3 | 0.98 |
| | | | | | | | |
| NN matching with replacement and caliper 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.246 | 134.69 | 0.000 | 19.8 | 14.7 | 105.2 | 0.12 |
| Matched | 0.026 | 11.12 | 0.802 | 7.3 | 5.7 | 38.4 | 1.39 |
| Radius with caliper 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.246 | 134.69 | 0.000 | 19.8 | 14.7 | 105.2 | 0.12 |
| Matched | 0.008 | 3.33 | 1.000 | 4.0 | 2.9 | 18.6 | 0.42 |
| **Kernel with bandwidth 0.06** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.246 | 134.69 | 0.000 | 19.8 | 14.7 | 105.2 | 0.12 |
| **Matched** | **0.005** | **2.19** | **1.000** | **2.9** | **2.8** | **13.7** | **0.16** |

Source: YL data: own computation

Appendix 5.3 summary result for the balancing test using various matching algorithms, for SNNP region

| NN matching without replacement | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LRchi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.150 | 48.43 | 0.000 | 21.3 | 20.4 | 102.9 | 0.30 |
| Matched | 0.036 | 4.73 | 0.997 | 7.1 | 8.5 | 44.6 | 0.77 |
| NN matching with replacement and caliper 0.05 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | Mean Bias | Med Bias | B | R |
| Unmatched | 0.150 | 48.43 | 0.000 | 21.3 | 20.4 | 102.9 | 0.30 |
| Matched | 0.026 | 3.34 | 1.000 | 6.0 | 4.3 | 37.6 | 0.94 |
| **Radius with caliper 0.05** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.150 | 48.43 | 0.000 | 21.3 | 20.4 | 102.9 | 0.30 |
| **Matched** | **0.009** | **1.11** | **1.000** | **3.2** | **2.8** | **17.4** | **0.11** |
| Kernel with bandwidth 0.06 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.150 | 48.43 | 0.000 | 21.3 | 20.4 | 102.9 | 0.30 |
| Matched | 0.029 | 3.79 | 1.000 | 6.9 | 6.6 | 33.7 | 0.16 |

Source: YL data: own computation

Appendix 5.4 summary result for the balancing test using various matching algorithms, for Tigray region

| 5-NN matching with a caliper of 0.05 | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.094 | 40.38 | 0.001 | 15.1 | 10.2 | 74.5 | 0.62 |
| Matched | 0.016 | 12.21 | 0.787 | 4.0 | 3.4 | 29.5 | 2.37 |
| **Radius with a caliper of 0.05** | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.094 | 40.38 | 0.001 | 15.1 | 10.2 | 74.5 | 0.62 |
| **Matched** | **0.015** | **11.75** | **0.815** | **3.7** | **3.6** | **29.0** | **2.51** |
| Kernel with bandwidth 0.06 | | | | | | | |
| Sample | Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B | R |
| Unmatched | 0.094 | 40.38 | 0.001 | 15.1 | 10.2 | 74.5 | 0.62 |
| Matched | 0.014 | 10.69 | 0.872 | 5.2 | 5.0 | 27.8 | 2.28 |

Source: YL data: own computation