Women Empowerment through Public Employment Programmes: Evidence from Ethiopia*

Master Thesis by Till Kadereit

Sciences Po Paris

Department of Economics

Supervisor: Prof. Benjamin Marx

Abstract

Empowering women represents one of the pillars for economic development, particularly in low-income settings. At the same time, Public Employment Programmes increasingly become a popular policy tool to promote growth and social security. Focusing on the Productive Safety Net Programme (PSNP) in Ethiopia, this study is one of the first to empirically evaluate the effects of a large-scale Public Employment Programme on the social and economic empowerment of women. Drawing on four cross-sections of the Demographic and Health Surveys (DHS) and using a Diff-in-Diff and triple differences approach, I find mixed results for programme effects on women empowerment. Whereas the PSNP did not positively affect many of the evaluated indicators, I find negative coefficients for early childbearing and early marriage throughout the analysis. Although they are insignificant and must be interpreted with caution, they could be driven by a significant relative increase in work of young women. Finally, while the overall work status of all women appears to be largely unaffected by the PSNP, there is some evidence that the programme crowded out self-employment and working for family members among women in rural Ethiopia.

*Acknowledgements: First and foremost, I would like to thank my supervisor Prof. Benjamin Marx for his great guidance, inspirations and support throughout the entire process of writing this thesis. I am also grateful to Jules Gazeaud for providing me with an additional data source, and to many others for giving helpful advises in the work with data. Finally, I wish to extend my gratitude to Laura, Martin, Marvin, Mathilde, Tim, Valentin and my parents for the interesting discussions, tips and their valuable support. All errors are my own.

Contents

List of Tables

List of Figures

1 Introduction

Being part of a broader employment and social protection policy, Public Employment Programmes (PEP) have become a commonly used policy tool to provide income security and increase social protection, particularly in developing economies. Targeting the most vulnerable groups in society, they serve as an instrument to tackle un- and underemployment, improve participants' employability through work experience and increase consumption through additional income (ILO 2021). The worldwide largest workfare programme is the Indian Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS), which has provided 2.3 billion person-days of employment in 2010-2011 alone and shows that these interventions can take on large dimensions and enjoy prioritization by policymakers (Imbert and Papp 2015). Less attention has been paid to similar employment programmes in Africa, where the Ethiopian Productive Safety Net Programme (PSNP) is one of the largest Social Protection Programmes. Recently, it addressed 10 million people and it has already been regarded as a model for the continent (Berhane et al. 2014; Lavers 2019). However, the question whether these policy interventions can empower women in developing countries has empirically received little attention so far. Particularly in Sub-Saharan Africa, women are often trapped in a low-empowerment equilibrium, where limited control over their body and early childbearing further limit their low labor market participation and increase the dependence on men (Bandiera et al. 2020). Many women in this region work in insecure and poorly paid jobs that give little hope of advancement and whereas fertility levels are mainly decreasing around the world, least developed countries in Sub-Saharan Africa are still characterized by high fertility levels (UN Women 2021; UN 2017).

This thesis aims to fill the gap in the literature and empirically evaluates the Ethiopian Productive Safety Net Programme (PSNP) in its impact on the empowerment of women, using four repeated cross-sections of the Demographic and Health Surveys (DHS) and a Diff-in-Diff and triple differences approach. The paper is guided by the question, whether the PSNP as a large-scale Public Employment Programme was effective in strengthening the role of women in a low-income setting in Sub-Saharan Africa and improved their situation in the long term along different social and economic outcomes. To answer this question, I compare changes in indicators of women empowerment between districts where the PSNP has been implemented and other districts in Ethiopia over time.

Although the topic has not been a major focus in the literature so far, several aspects point to the importance of evaluating Public Employment Programmes in terms of possible effects on women's role and welfare.
From an economic perspective, women empowerment is considered to have the potential to spur growth and development. Not least due to the belief that discrimination against women hinders development, policymakers put women empowerment among the top priorities on the agenda. For instance, evidence has shown that less women rights caused inefficiencies, such as too little investment or lower productivity on farms in low-income settings, and at the same time it has been shown that empowering women can improve aspects of children's welfare, like health or nutrition (Duflo 2012). What is more, given evidence

that women have generally lower fertility preferences than men, empowering women and increasing their agency position within the household can play an important role in curbing population growth, which is positively correlated with poverty (Banerjee and Duflo 2011).
In this respect, social protection schemes and safety nets are considered to have a great potential in terms of decreasing fertility, fighting poverty and potentially empowering women (Banerjee and Duflo 2011). More specifically, when they are well-designed, Public Employment Programmes can play an important role in promoting gender equality and pro-poor development. Among other things, they can lead to better women and men participation ratios through the provision of equitable access to jobs, generate income and provide food security. Through the creation of assets and infrastructure that enhance access to communal resource inputs and basic social services, they can also address women's burden of unpaid work (Antonopoulos 2007). Also, having less labor market prospectives can result in unequal treatment of female members of the household, for instance because they are considered to need less food or education. Providing labor market opportunities can therefore be a catalyst for improving women's and girls' position in the household (Duflo 2012)
As I will point out in the following chapter, the PSNP has a particular focus on social protection and also includes several gender-specific provisions (McCord and Paul 2019; Holmes and Jones 2011). There can thus be reasonable expectations that the PSNP can help to break women's low-equilibrium trap of limited labor market participation, low human capital levels and high fertility, and can help to promote gender equality and growth in Ethiopia. Providing labor market opportunities and income could not only strengthen women economically, but also go hand in hand with an improved role in the household, agency power and social empowerment.
Thus, my thesis can help to find out which mechanisms and policies work to promote women empowerment in general, and whether Social Protection and Employment Programmes in particular can be a response to the ongoing discrimination of women.

The remainder of this paper is organized as follows. The subsequent part gives a brief overview of how the thesis contributes to the literature, while chapter 2 provides background information on the programme. Moreover, the chapter introduces a brief conceptual framework that will guide the empirical proceeding. Chapter 3 and 4 present the empirical strategy and the data used in the analysis. Afterwards, the main results of the thesis as well as different robustness checks and a brief discussion of the findings are presented in chapter 5. Finally, I draw the central conclusions and provide an outlook for potential future research in this field.

1.1 Literature

Evaluating Ethiopia's Productive Safety Net Programme in terms of women empowerment mainly contributes to three different strands of literature. First, empirical studies that evaluate Public Employment Programmes in its impact on the empowerment of women remain limited so far. To the best of my knowl-

edge, the only empirical evaluation of the PSNP in this respect was carried out by Gelegay and Lecoutere (2019). The authors examine the effect of the conditional cash transfer (i.e. the public works) component of the PSNP on different aspects of women's economic empowerment in the Tigray region. Thereby, they rely on Propensity Score Matching methods and cross-sectional household data from 2015 and measure economic empowerment through 10 indicators along the five key domains of the Women's Empowerment Agricultural Index, which are production, resources, income, leadership and time. Running their specifications separately for women- and men-headed households, Gelegay/Lecoutere find differential impacts of the PSNP on the outcomes, as well as heterogeneity in effects for women and men-headed households. The only significant positive finding suggests that public works participation improves women's access to- and decision-making about credit, but on the other hand there are negative effects on decision-making power over agricultural production and the transfer of assets among women-headed households (Gelegay and Lecoutere 2019). My thesis adds to this work in an important way, extending the setting to other regions than Tigray and relying on a Diff-in-Diff approach, which allows me to exploit trends over time. Also, I put a stronger focus on outcomes of women's control over their body, thus shedding more light on the role of the PSNP for social empowerment.
In another setting of Public Employment Programmes, Amaral et al. (2015) evaluate India's MGNREGS in terms of impacts on gender-based violence using a differences-in-differences approach. Their findings suggest that increased female labor market participation following the MGNREGS increased gender-based violence like kidnappings and domestic violence, whereas dowry deaths decreased (Amaral et al. 2015).

My thesis also contributes to previous impact evaluations of Public Employment Programmes on outcomes others than women empowerment, particularly of the PSNP in Ethiopia. Hoddinott and Mekasha (2020) evaluate the effects of the PSNP as social protection scheme on household size and factors that determine household size, using a Diff-in-Diff approach and four survey rounds from 2006 to 2012. Interestingly, they find that household size increased slightly, but they also provide some evidence that the fertility rate decreased among women who were participating in the PSNP. The authors conclude that the increase in household size was mainly caused by an increase of adolescent females in the household due to reduced out-migration, which implies that the PSNP caused a delay in marrying out young women (Hoddinott and Mekasha 2020). In one of the most influential studies in the context of the PSNP, Gilligan et al. (2009) assess the programme on a range of outcomes like food security, consumption levels, asset holdings and investments through better credit access, 18 months after its implementation. Using cross-sectional data and propensity score matching methods, the authors find that impacts critically depend on how programme participation is defined: Whereas payments through the Public Works component alone seem to have little impact, Gilligan et al. find significantly positive effects on food security, productive borrowing and the use of agricultural technologies when participants received PSNP payments combined with productivity enhancing transfers from the OFSP (Other Food Security Programme). Also, they do not find evidence that PSNP transfers reduced labor supply of beneficiaries to wage employment (Gilligan et al. 2009). Gazeaud and Stéphane (2020) assess the productive effects of infrastructure generated

through Public Works in the PSNP and thereby take differences in the intensity of PSNP coverage across districts into account. Relying on georeferenced administrative and high-resolution satellite data, they find no evidence that Public Works had a significant impact on agricultural productivity (Gazeaud and Stéphane 2020). In terms of labor market impacts, Schinaia (2016) finds no evidence for effects of the PSNP on overall employment in rural Ethiopia, and neither on private sector labourer's wages, whereas there is evidence for an increase in non-agricultural self-employment (Schinaia 2016). In India, Imbert and Papp (2015) examine the labor market implications of the large-scale Public Employment Programme MGNREGS. Comparing trends in early and late programme implementation districts, they provide evidence for a crowding out of private sector work and increased private sector wages, as well as for an increase in public sector work. Imbert/Papp also find significant redistributive effects to the poor through increased equilibrium wages, but point out that the fiscal costs of the programme were higher than the estimated welfare gains for the poor (Imbert and Papp 2015).

Finally, considering the PSNP as a potential policy tool to empower women, my work adds to the body of literature that explores policy interventions which aim at promoting women empowerment and improving women's welfare in low- or middle-income settings. Bandiera et al. (2020) use a RCT in a setting with a very young population in Uganda, to investigate the effects of a two-pronged intervention that simultaneously provides life-skills and vocational training to adolescent girls. The intervention aims to break a vicious circle: On the one hand, restricted labor market access and low human capital investment cause early marriage and early childbearing among young women. On the other hand, this further deteriorates their future labor market prospects, resulting in a low-equilibrium trap. The results show strong and lasting positive effects of the RCT on a range of outcomes like fertility, early marriage, violations and self-employment (Bandiera et al. 2020). Given the nature of the PSNP design, my paper can also contribute to the vast literature that evaluates Conditional Cash Transfer (CCT) programmes. In the context of a RCT in Malawi, Baird et al. (2016) emphasize the important role cash transfers can play in improving lives of young women in Sub-Saharan Africa. The Zomba Cash Transfer Program provided monthly cash transfers for adolescent girls over two academic years, including an unconditional transfer and a treatment arm that provided transfers conditional on girl's school attendance. Significant positive effects on a range of outcomes like access to financial resources, improved health and schooling outcomes, decreased pregnancies and early marriages and improved agency suggest that such programmes can be an important policy tool to empower women in this setting (Baird et al. 2016). Many CCTs prescribe that household transfers are made exclusively to women, hoping to improve children's welfare and women's power within the household. Investigating the decision-making power and agency position of women in this context, Brauw et al. (2014) provide evidence that the worldwide largest CCT programme, the Bolsa Família in Brazil, increased women's decision-making power in the household in aspects like contraceptive decisions or children's health expenses, whereby it appeared to be effective only in urban areas (Brauw et al. 2014). Finally, Stecklov et al. (2007) investigate effects on childbearing for three randomized CCTs in Latin America that make transfers conditional on children's school attendance and family member's

health care, and show that the programme design matters to avoid unintended effects. Whereas the authors find no impacts on fertility for the PROGRESA (Mexico) and RPS (Nicaragua), significant increases in fertility through the PRAF in Honduras are apparently caused by inadvertently created incentives for childbearing due to the programme design, which allowed parents to obtain increased benefits by bearing children after the programme start (Stecklov et al. 2007).

This thesis shall enrich the mentioned strands of literature, providing novel evidence on the potential of Public Employment Programmes and CCTs to empower women. Whereas the chief focus of my work lies on the social dimension of women empowerment, an aspect that has hardly been investigated in the context of Public Employment Programmes so far, this paper provides a comprehensive evaluation of several of the most important factors in empowering women and promoting development.

2 Background

2.1 The Productive Safety Net Programme (PSNP)

The Ethiopian Productive Safety Net Programme (PSNP) is one of the largest Social Protection Programmes in Africa and was introduced in 2005 as central component of the Ethiopian government's food security strategy, first addressing approximately 5 million rural inhabitants. Until then, responses to food insecurity had mainly been dominated by emergency food aid in Ethiopia, which had proven inefficient in protecting livelihoods in the long run. The PSNP was designed to address both chronical and transitory food insecurity and aimed to break Ethiopia's chronic dependency on food aid. Its main objectives include helping households to smooth consumption, protect assets and strengthen their resilience to shocks and creating community assets that could increase household productivity. Initially, the PSNP was linked with the second component of the food security strategy, the OFSP (Other Food Security Programme), which aimed to build household assets through the provision of fertilizer or credit, for instance (Devereux et al. 2008; World Bank 2012; Holmes and Jones 2011). The PSNP is an international flagship program in terms of both scope and its partnership approach, formed in 2003 by a broad coalition of Government, UN agencies, donors and civil society. In 2009, the annual budget was $360 million and thus about 1.2% of the Ethiopian GDP (World Bank 2012).
Particularly in Sub-Saharan Africa, Public Employment Programmes have increasingly become linked with the interest in providing productive safety nets and transformative social protection, extending the public employment provision through rural development activities. The PSNP in Ethiopia is considered a leading example of this productive safety net model of PEPs and thus has a particular focus on social protection (McCord and Paul 2019).

2.1.1 Roll-Out and Targeting Process

The Productive Safety Net Programme was planned to be launched in January 2005 and finally started with a slight implementation delay in April 2005 (Sharp et al. 2006). So far, the programme comprises four

different phases. The first and second phase were running from 2005-2006 and 2007-2009, respectively, followed by a third phase that brought further extensions between 2010 and 2015. Phase 4 of the PSNP was launched in January 2015 and is designed to eventually scale up to a national programme with an annual caseload of about 10 Mio. clients (World Bank 2012; MOAD 2014). Ethiopia is administratively divided into 9 regions and two autonomous cities. On the third administrative level it is divided into 710 Woredas, which are equivalent to districts, followed by Kebeles, the lowest level of elected government in Ethiopia. Starting in 192 Woredas and addressing about 4.5 Mio. households in 2005, the PSNP expanded to 290 Woredas during the first phases. In 2006, the pastoral region Afar joined the program, first of all only under the Direct Support component, the Somali region joined in 2008. However, whereas the increase from 192 to 234 Woredas from 2005 to 2006 could be attributed to an increased coverage of the PSNP, a further increase of 234 to 290 Woredas was almost entirely driven by an administrative reform and the resulting division of Woredas. In 2009, the programme addressed approximately 7.6 Mio. beneficiaries in 290 Woredas and 8 regions, thus covering about 10% of the country's population and 40% of all Woredas (World Bank 2012; World Bank 2010).

Importantly, the selection of PSNP beneficiaries is subject to a general targeting procedure, which is a combination of geographic targeting and a community-based selection process that aims to identify particularly food-insecure households. Beneficiary selection takes place at Woreda, Kebele and community level. Using the existing emergency response system, the government first identified areas that were food insecure, namely the regions and Woredas that had received food aid for the preceding three years or longer. Having targeted the food insecure districts, Woreda administrators select chronically food insecure Kebeles within a Woreda, assigning PSNP quotas. Finally, Food Security Task Forces are in charge of identifying eligible participants at community-level. The Program Implementation Manual manifested criteria for eligible chronically food-insecure households, which must have faced continuous food shortages in the preceding three years and received food assistance prior to the PSNP, have suddenly become more vulnerable and not be able to support themselves anymore, or be without family support or any other social protection (World Bank 2012).

2.1.2 Components

Providing food or cash transfers to chronically food insecure households for several months per year and up to five years, the Safety Net aims to make households food sufficient, such that they will finally graduate from the programme (Devereux et al. 2008). Two central components constitute the PSNP: The Public Works (PW) and Direct Support (DS) component, whereby most of the benefits are conditioned on participating in public works.

Under the PW component, all chronically food insecure households that are able to provide labor force, receive transfers for participating in public works. Public work is offered six months per year and carried out from January to June during the agricultural slack season. It comprises mainly small-scale, labor-intensive community projects that provide unskilled employment. To receive the full monthly payment, participants must work a minimum number of days, which is determined by family size and the number of able-bodied

adults in the household. In general, this is set as 5 days per household member per month. Moreover, participants worked on average 8 hours per day. The programme design envisages that public works complements transfers by developing rural productivity and infrastructure through community assets. Eligible for Public Works are all labor-bodied women and men of at least 18 years who are targeted by the PSNP, whereby the age limit explicitly aimed to avoid early school drop-out of young participants (World Bank 2012; Sharp et al. 2006). During the design phase of the PSNP concerns arose that the programme could exclude most vulnerable groups, particularly labor poor households. Therefore, the Direct Support (DS) component was introduced to complement the PW arm. It guarantees unconditional transfers to programme participants who are not able to participate in public works. This includes people with disabilities or pregnant women, for instance.
First estimations of the distribution of participants reported that approximately 80% of PSNP beneficiaries participate under the Public Works component and 20% receive unconditional direct support transfers. Transfers, provided in cash, food or a mix of both, are the same for public works and direct support clients and are made on a monthly basis for six consecutive months (World Bank 2012).

2.1.3 Gender Perspectives in the PSNP

When the Ethiopian government and international donors designed the programme in the early 2000s, they acknowledged the need to include gender aspects in its guidelines to respond to the vulnerabilities of women in rural Ethiopia. Thus, the PSNP generally has a strong focus on women's role in agriculture and addresses women's vulnerabilities on different levels (Holmes and Jones 2011; World Bank 2012). More precisely, gender-provisions were incorporated along four categories in the initial design of the Productive Safety Net Programme. First, the PSNP includes analyses of gender-specific vulnerabilities that women face in terms of socio-cultural gender roles and life-cycle factors and accounts for women's reproductive labor. For instance, it shall be guaranteed that pregnant women are shifted from the Public Works to the Direct Support component. Moreover, the PSNP acknowledges the importance of women's participation in public works through provisions to include female-headed households and to promote women's participation in community decision-making. Thirdly, community assets created through the programme are previsied to tackle gender issues. Particularly the provision of assets like fuelwood and community water sources is designed to address women's time poverty. Finally, the PSNP puts the inclusion of women in committee structures at state and Woreda level on the agenda (Holmes and Jones 2011; World Bank 2012).

In terms of concrete policies, specific approaches that have been developed to facilitate women's participation and address gender inequality include the institutionalization of explicit quotas for female and female-headed household participants, the guarantee of equal wages for women and men and the support of female participants for better access to credit. Moreover, the PSNP design allowed for time-off during pregnancy and flexibility in women's working hours to balance domestic and care work responsibilities, and it envisaged the involvement of women in decision-making processes regarding the types of community assets built under the PW (Holmes and Jones 2011). First fieldwork evidence in some Woredas reported

positive implementation experiences in this respect, for instance that women accounted for approximately half of all participants, and that women indicated to have improved their standing in the community and being accounted more respect by their husbands due to their participation in PSNP projects (Sharp et al. 2006; Holmes and Jones 2011; World Bank 2012).

2.2 Conceptual Framework

2.2.1 Women Empowerment as a Concept

Since at least after Amartya Sen's article "More than 100 Million Women are Missing" (Sen 1990), the awareness that women are discriminated against in many aspects worldwide was more deeply rooted in politics and the economic literature. It seems to be consensus that the grievance of gender inequality, which exists in various forms in developing and developed countries, must be corrected by empowering women, for reasons of equity as well as efficiency and development. But what exactly is women empowerment and how can it be measured? Before I will empirically evaluate a policy tool that potentially helps to strengthen the role of women, it is all the more important to discuss what will eventually be evaluated and how a Public Employment Programme like the PSNP could make a difference in this respect.

Empowerment as such is a broad concept and its measurement often turns out to be complicated – some even doubt that it can be clearly defined or measured. In a highly influential paper, Kabeer (1999) defines empowerment as "the processes by which those who have been denied the ability to make choices acquire such an ability" (Kabeer 1999, p. 437). She distinguishes between three interlinked dimensions of women empowerment: Resources, agency and achievements. Resources do not only relate to material, but also to human and social resources that serve to enhance the ability to make choices. Frequently used as indicator in the resource dimension is the concept of control, such as having the final say in important decisions. Agency is often put equal to decision-making but can also take forms of bargaining or resistance. Examples to measure agency include the participation in public action or the incidence of male violence. As frequently used indicators to measure achievements, Kabeer lists outcomes like owned assets, female education or working before marriage (Kabeer 1999).
Duflo (2012) defines women empowerment as "improving the ability of women to access the constituents of development" (Duflo 2012, p. 3), which include health, education, earning opportunities, legal rights and political participation and representation. She identifies deficits in gender equality particularly along these dimensions and mentions women's time poverty as another central aspect of inequality between women and men. In both rich and poor countries women spend considerably more time for housework and care responsibilities, which leaves less possibilities for market participation and an autonomous income (Duflo 2012). Bandiera et al. (2020) report three key dimensions of women empowerment, which are political and economic empowerment as well as control over the body. In the context of their RCT in Uganda, they distinguish between economic empowerment, social empowerment and future aspirations and attitudes of young women, whereby social empowerment is defined in its most basic form of having control over one's body. This concept addresses the prevalent negative interaction of low labor market opportunities

and limited control over the body of young women in the Sub-Saharan country (Bandiera et al. 2020). The mechanism could also be relevant to my setting, as Ethiopia continues to be characterized by a high fertility level, also in comparison to the least developed countries, with a total fertility rate of 4.59 in the recent past (UN 2017). For the rest of my analysis, I will broadly stick to this categorization of women empowerment, adding women's time poverty and nutritional status to my concept.
In the subsequent part, I will briefly discuss how I expect the PSNP to address different dimensions of women empowerment and potentially change women's situation in rural Ethiopia for the better. This shall guide the empirical proceeding and help to explain the obtained results.

2.2.2 Women Empowerment and the PSNP

One major focus of this study lies in assessing the Productive Safety Net Programme's impact on the social empowerment of women, reflected by the control over one's body. There are various channels through which the PSNP could affect outcomes in this dimension. Broadly speaking, development goes hand in hand with reduced fertility and, more specifically, fertility has been shown to decrease with income and education (Duflo 2012). Previous assessments have shown that conditional cash transfers (CCT) can have the potential to decrease fertility through increased income and improved human capital accumulation of beneficiaries' children, for instance (Todd et al. 2010). Strikingly, the PSNP could affect fertility and family planning through another channel. If female beneficiaries receive transfers conditional on participating in public works, the opportunity cost of childbearing and staying at home will necessarily be higher, which creates an incentive to delay or suspend pregnancy and keep on benefiting from the PW transfers (Hoddinott and Mekasha 2020). Thereby, it should still be kept in mind that the PSNP allowed for the possibility to be assigned to the Direct Support component in case of pregnancy. Nevertheless, work opportunities can provide a strong channel for social empowerment and the PSNP could thus have the potential to break the aforementioned vicious circle of high fertility levels and little human capital investment, which could then be reflected by lower fertility, increased use of contraceptives or delayed marriage.
Moreover, the Productive Safety Net Program can be expected to have a positive impact on women's gender perceptions and decision-making power in the household. Literature that assesses the impact of conditional cash transfer programmes on women's role in the household, provided evidence that cash transfers that were made exclusively to women increased their bargaining- and decision-making power (Brauw et al. 2014). If the PSNP was successful in providing work and cash transfers to women who were not employed before, this could change their role and bargaining power within the household, as well as their own perceptions of gender roles.

The PSNP comes in the form of a labor intensive cash-for-work programme and safety net and could therefore importantly affect women's economic situation. Given that the large share of PSNP beneficiaries participates in the PW component and receives transfers conditional on working in community projects, it is of particular interest to examine women's work status and evaluate, if the Safety Net was successful

in integrating women into the programme.
Secondly, I evaluate educational outcomes for cohorts of young women. Expected effects in terms of education are ambiguous. On the one hand, the PSNP could positively affect families' education decisions towards girls if it can promote economic development and food security. Additional income received through the PSNP could make households less dependent on girls' labor force or make them invest some of the received cash transfers in their daughters' education. On the other hand, there can be concerns that the public works component induces early school drop-out by providing incentives to participate in the programme, although the PSNP design officially aimed to rule out any inference with children's education.
Finally, I investigate impacts on women's time poverty in my analysis. In developing countries, particularly water insecurity is a heavily gendered driver of poverty. Women bear most of the responsibility to collect water and are subject to physical and psychological strain, an increased risk of gender-based violence, and they loose valuable time that could be spent on other productive activities like schooling or employment. Social protection programmes have the potential to promote improvements in water security, among other things through public works projects that enhance the water infrastructure (Lowe et al. 2019). As the PSNP aims to approach the created community assets through a gender-sensitive lens and designed public works activities also to create community water sources, I will measure programme impacts on women's time poverty by the time women needed to get to the next water source and back (Holmes and Jones 2011).

3 Empirical Strategy

To measure impacts of the Ethiopian Productive Safety Net Programme on women empowerment, I rely on a differences-in-differences approach and repeated cross-sectional data. My identification strategy consists in comparing changes in outcomes of women empowerment between Woredas that were targeted by the PSNP and non-targeted Woredas over time. A major issue in a simple cross-sectional comparison would be the targeting process. Given that the PSNP did not follow random assignment but instead explicitly targeted food insecure Woredas that relied on food aid in the preceding years, unobserved differences between treatment and control group that are correlated with women empowerment outcomes are likely to occur, causing endogeneity issues. It is reasonable to assume that I would underestimate programme effects when treatment Woredas are poorer and thus also likely to have worse initial women empowerment outcomes. A Diff-in-Diff strategy as quasi-experimental approach can account for permanent differences between treatment and control districts as well as general time trends and thus provides a good opportunity to evaluate PSNP impacts on women empowerment (Imbens and Wooldridge 2007).

A threat to validity in a standard differences-in-difference approach occurs when the outcomes of interest would not have followed the same trends in treatment and control Woredas in absence of the PSNP programme (Imbert and Papp 2015). For instance, bias could result from time-varying group-specific shocks that may be correlated with women empowerment outcomes but are not captured by the

time-invariant district fixed effects. To address the possibility of differential trends between treatment and control Woredas, I first include a set of time-varying and time-invariant district-level controls. In Ethiopia, one example of time-varying shocks that could cause differential trends are climate hazards. The country is particularly vulnerable to climate hazards and is regularly hit by droughts. Droughts have thereby shown spatial patterns in the country and affected some regions more frequently, whereby the North Eastern, Eastern and South Eastern lowlands are significantly drier and constitute the most vulnerable areas in Ethiopia. Moreover, droughts have been shown to be linked with social and economic disruptions and are thus likely to affect many of the women empowerment outcomes evaluated in this thesis, particularly given that water-related shocks tend to affect women more severely (Alem Mera 2018; Lowe et al. 2019). Figure 1 shows that the sampled Woredas in my panel are geographically dispersed, whereby the treatment group is mainly located in the dry northeastern part of the country. This could point to the fact that both groups are indeed not exposed to the same (rainfall) shocks. I therefore include the percentage deviation of an annual 5-year average in rainfall from the long-term mean and its square as time-varying district-level controls. Secondly, I include pre-treatment measures of the population density, female literacy rate, a self-constructed poverty rate and an infrastructure index, which are interacted with a post-treatment dummy. Whereas rainfall serves as indicator for the incidence of productivity shocks and thus also for the general economic development of the past years, the time-invariant control variables are included to account for potential differential trends that could be caused by initial differences in these variables between the groups. Finally, a possible violation of the common trend assumption is further addressed by the use of triple interactions for some of my outcomes, as explained in the subsequent section.
The central regression equation of the analysis then looks as follows:

$$y_{ijt} = \beta(Post_t \times PSNP_j) + \delta X'_{jt} + \eta(X'_{jt} \times PSNP_j) + \mu(Z'_j \times Post_t) + \alpha A'_{ijt} + \gamma_j + \lambda_t + \epsilon_{ijt}$$

In this equation, y_{ijt} is the outcome of woman i, observed in Woreda j at time t. *PSNP* is a dummy that equals one if a Woreda was targeted by the PSNP and *Post* is an indicator for one of the post-treatment years 2011 and 2016. β_{jt} is the coefficient of interest: It shows whether the PSNP improved women empowerment outcomes in targeted Woredas, relative to the control group. Furthermore, time fixed effects λ_t are included to control for general time trends that affect all Woredas, and Woreda fixed effects γ_j control for time-invariant Woreda-specific characteristics. X'_{jt} is a vector of the above-mentioned time-varying controls and Z'_j a vector of the time-invariant pre-treatment characteristics at Woreda-level, which is interacted with the dummy for the post period. To account for possibly differential effects of rainfall variations between the control and the much drier treatment Woredas, I additionally interact the rainfall variables with the dummy for treatment status, as Imbert and Papp (2015) do in a similar setting. Finally, to increase the precision of the estimates, I also include several individual-level controls in vector A'_{ijt}, namely age-groups, religion, sex of the household head and access to electricity in the respondent's household.

Most of the observed women empowerment outcomes are binary variables, like being married or having

given birth in the past three years. All binary outcomes in this analysis are hence estimated in a Linear Probability Model (LPM), using OLS estimations. Programme impacts are then interpreted as changes in the probability that these outcomes occur. In the robustness checks, I will additionally investigate programme effects using a Probit model for the binary dependent variables.
Since standard errors computed in a conventional Diff-in-Diff are likely to be serially correlated, they would understate standard deviations of treatment effects and thus overestimate the t-statistics (Bertrand et al. 2004). Therefore, standard errors are clustered at Woreda-level, the level of treatment assignment, throughout the analysis to allow for autocorrelation within Woredas over time.

3.1 Measuring Early Childbearing, Early Marriage and Education: DDD

To measure possible effects of the PSNP on early marriage, early childbearing and education, I slightly adjust my empirical strategy. For these outcomes, it is of interest to examine whether there are effects on a certain age cohort. Aiming to observe whether the PSNP prevents women from marrying or bearing children in early ages, the sample must be restricted to young women. The same idea applies to education: The PSNP could only have affected educational attainment, when girls were still in school age when the programme started to run.
One could thus restrict the sample to a certain age limit and run the same Diff-in-Diff regression as before. However, there are advantages of running a triple differences specification here. Whereas the Diff-in-Diff model removes effects of time and place, a triple difference approach can reduce the bias in policy evaluations when the outcome is not only determined by policy, place and time, but additionally by another variable (Berck and Villas-Boas 2016). Thus, I interact the dummies for program implementation and the post period with a third dummy, which equals one if participants are of the young cohort. This allows to compare changes in fertility, marriage and educational outcomes of young women between treatment and control Woredas, but also between young and older women within both groups, which is supposed to be more robust. On the one hand, the analysis controls for trends in outcomes that are common to young women across Woredas unrelated to the PSNP, by comparing changes between young women in treatment and control Woredas. On the other hand, by comparing changes in outcomes between the age groups, a triple differences approach can also account for general differential trends in outcomes between the treatment and control group (Imbens and Wooldridge 2007). Thus, although one should recall that in this setting the treatment was not exclusively aimed at the young group, this proceeding can at least partly address general group-specific (differential) trends in outcomes.
In this case, the regression equation from before becomes as follows:

$$\begin{aligned} y_{iajt} =& \beta(Post_t \times PSNP_j \times Young_a) + \delta X'_{jt} + \eta(X'_{jt} \times PSNP_j) + \gamma(Z'_j \times Post_t) \\ &+ \nu(Z'_j \times Post_t \times Young_a) + \alpha A'_{iajt} + \gamma_{jt} + \lambda_{at} + \theta_{aj} + \epsilon_{iajt} \end{aligned}$$

Where now γ_{jt}, λ_{at} and θ_{aj} represent Woreda and time fixed effects, but also fixed effects for the young cohort, as well as all respective double interactions between the three fixed effects. Moreover, the time-

invariant Woreda-level controls are additionally interacted with the dummy for the young cohort. Outcome Y_{iajt} is observed for individual i of age a in Woreda j at time t. The coefficient of the triple interaction β shows whether the PSNP was successful in preventing young women from engaging in early marriage and childbearing or in increasing young women's educational attainment, thus potentially strengthening their future labor market trajectories.

Finally, it remains to specify what I understand as young cohort. Particularly for marriage and education, the definition must make sure that any observed effect for the young cohort can be attributed to the PSNP. Thus, individuals must be of an age that made them a candidate for early childbearing, early marriage or receiving education during the PSNP implementation period. For early marriage and childbearing, I define the young group as women between 15 and 22. Similarly, I define the young cohort for education as women between 15 and 20, thus guaranteeing that young-group respondents were still of school age throughout the period of interest and thus allowing the PSNP to have effects on the years of schooling.

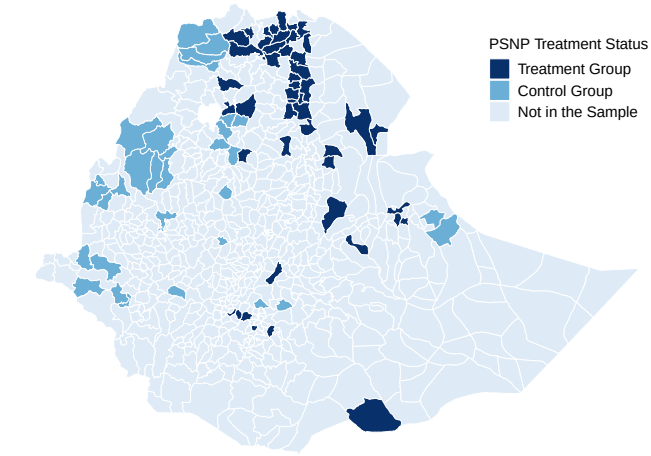
3.2 Treatment Assignment

As explained in the second chapter, the PSNP was not implemented in all Woredas at the same time, but rather followed a continuous roll-out. After having started in 192 Woredas in 2005, the programme expanded to 290 Woredas until 2009 (World Bank 2012). My data sources for the treatment assignment are from 2007 and 2009[1] and can thus account for the increase in programme coverage. Given the availability of treatment assignment data, I define the treatment group as all Woredas that have been targeted by the PSNP in either 2007 or 2009. The PSNP programme design intended to support households with transfers up to a period of five years (Devereux et al. 2008). It is thus likely that Woredas were also targeted to participate in the long term and that most of the Woredas that have been targeted in 2007 or 2009 were initially assigned to the PSNP at the time of the programme start. Also, it is reasonable to assume that the PSNP could have made a difference even if a Woreda was not exposed to the programme during the entire assessment period. Figure 1 shows three maps of Ethiopia at the Woreda-level. The first map shows the geographic distribution of the targeted and non-targeted Woredas in my panel, the second map adds treatment intensity levels as percentages of the targeted population within the treatment group, and the third map illustrates the general geographical targeting distribution of the PSNP in Ethiopia. Among other things, the maps reflect the programme's strong focus on the drier North- and South Eastern parts of the country and show that the treatment and control group of my panel are geographically dispersed across the North-West and North-East of Ethiopia.

In general, my treatment definition does not distinguish between the two different components of the PSNP, Public Works (PW) and Direct Support (DS). But as mentioned in the previous chapter, the two program arms were intended to enter simultaneously. One exception is the region Afar, which entered the program in 2006 first only under the Direct Support component. However, the PSNP expanded to Afar and Somali entirely in Phase 3, which was running from 2010 to 2015 (World Bank 2012; MOAD 2014).

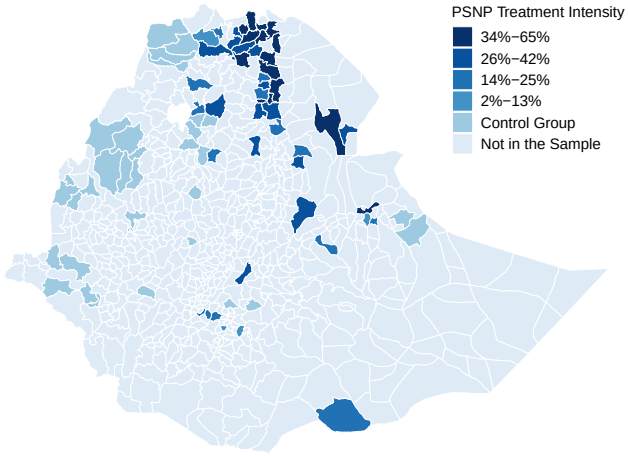
[1]Details on treatment assignment and the data sources are provided in the Appendix.

Figure 1: PSNP Targeting and Roll-out



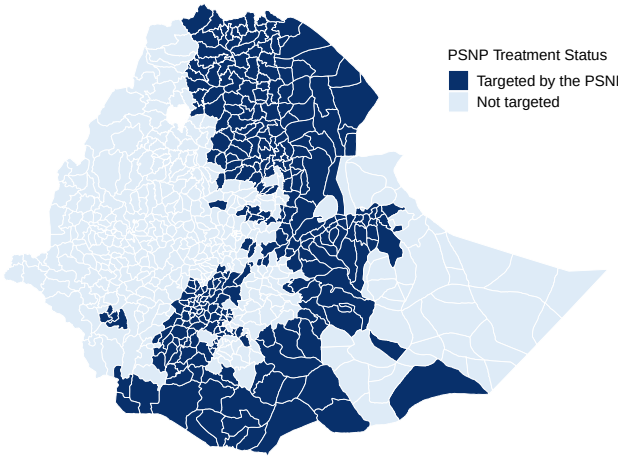
Source: Author's elaboration based on UNOCHA (2007) and World Bank (2010)

(a) Treatment status in the balanced Woreda–year panel



Source: Author's elaboration based on UNOCHA (2007), World Bank (2010) and Gazeaud/Stéphane (2020)

(b) Treatment Intensity in the balanced Woreda–year panel



Source: Author's elaboration based on UNOCHA (2007) and World Bank (2010)

(c) Treatment assignment of the PSNP in Ethiopia

A possible issue with having data for the treatment assignment of 2009 as latest point in time is that the PSNP could have expanded to more Woredas in the time between 2009 and 2011 or 2016, the years that define my post-programme sample. However, this appears to concern only a small number of Woredas, given that the number of treated Woredas was reported to be 319 between 2010 and 2015 (World Bank 2016). Thus, it is unlikely that some Woredas of my control group were assigned to the PSNP after 2009 and thus caused biases.

4 Data

This thesis relies on several data sources, whereby the Demographic and Health Surveys (DHS) (CSA 2000-2016) program represents the primary data source. Data for treatment assignment comes from the World Bank (2010) and UNOCHA (2007), data sources for the control variables include the US Climate Hazard Center and the NASA Socioeconomic Data and Applications Center.
The Demographic and Health Surveys are nationally representative population-based surveys and the DHS program has so far conducted more than 300 surveys in over 90 countries. Data sets are provided for different entities of the household, primarily for women. Recently, the DHS also provides additional data like AIDS Indicator surveys and geographical data (ICF 2021b). I use all four survey rounds that are available for Ethiopia, 2000, 2005, 2011 and 2016, as well as the geographic data sets for each year. The DHS surveys are not panel data, however, pseudo-panels can be constructed from repeated cross-sectional data when genuine panel data is not available, by grouping individuals that share common characteristics into cohorts (Verbeek 2008). Thus, I construct a balanced Woreda-year panel using the four cross-sections of the DHS survey. Since the DHS program only samples a subgroup of all Woredas per round that are not necessarily the same each survey year, the balanced sample finally comprises 78 Woredas. Whereas treatment takes place at Woreda-level, I measure outcomes at the individual-level and run the analysis on the micro-data.

Demographic and Health Surveys follow a two-stage probability sampling design and are stratified by geographic region and urban/rural area within each region. In the first stage, the Primary Sampling Units (PSU), which form the survey cluster, are selected with probability proportional to size (PPS) within each stratum. In the second stage, a fixed number of households is randomly selected within a cluster, following a household listing. Most surveys use about 25-30 households per cluster. To expand the number of available cases for certain areas or subgroups, most samples are selected with unequal probabilities. The surveys are carried out using four different questionnaires: Households, Men's, Biomarkers and Women's questionnaires. For my analysis, I only rely on the women's questionnaire and the respective data sets. Eligible for the interviews and thus included in the women's data set are all women aged 15 to 49 (Croft et al. 2018).
DHS samples are representative at the regional level as lowest administrative unit and any geographic information is not released below the second-level units (Zones). Since my identification strategy consists in comparing changes in Woredas (i.e. districts) over time, I additionally rely on the DHS geographic data set for each year. The DHS program georeferences survey cluster locations and makes them available in separate GPS coordinate data sets. Using GIS software, the geographic data sets allow to identify DHS observations at Woreda-level[2]. One issue with the geographic data sets is that survey clusters are geomasked to guarantee the respondents' confidentiality. Urban clusters are displaced a distance between 0 and 2 kilometres and rural clusters a distance between 0 and 5 kilometres, whereby a further 1% of rural

[2]Details on the construction of the panel and the identification of observations at Woreda-level are provided in the Appendix

clusters are displaced up to 10km (Burgert et al. 2013). However, this concerns only a few kilometres, and many of the treatment or control Woredas are neighbouring Woredas of the same group, such that the displacement of GPS clusters should remain a minor issue.

Since the Ethiopian Productive Safety Net Programme was exclusively aimed at the food-insecure rural population in the first phases[3], I restrict the sample to rural women and exclude all respondents living in urban areas throughout the analysis to avoid bias, as it was done in similar settings (Imbert and Papp 2015). Moreover, I drop all respondents from the sample that were not permanent residents of the household where they gave the interview.

4.1 Variables for the Analysis

Data for the outcomes of interest in this analysis is exclusively taken from the four DHS survey rounds. The concept and measurement of women empowerment that is used in this thesis is on the one hand based on the relevant literature, and on the other hand on the specific PSNP design and expectations about its effects, as discussed in chapter 2. In the social empowerment dimension, fertility is measured by the incidence of having given birth in the three years preceding the interview. Looking at the total number of births or at whether a woman has a child or not, may not be informative for identifying the short-term impact of the PSNP on fertility. Also, as Stecklov et al. (2007) point out, measuring childbearing should account for a likely occurring delay between changes in demand for children and conception.

My choice of a brief period of three years shall guarantee that most Woredas of the treatment group had already been covered by the PSNP at the time of conception and in the months before. Likewise, the same variable is used to measure early childbearing in a triple differences framework. Early marriage is also measured by a triple differences approach, using a dummy variable for whether the respondent was married or cohabiting at the time of the interview. Furthermore, I measure women's attitude towards sexual violence and control over the body by a dummy for whether a woman agrees that wife beating by the husband is justified when a wife refuses to have sex. Finally, to measure the use of contraceptives, I take a dummy that equals one if a respondent was currently using any modern or traditional contraceptive methods, whereby I restrict the sample to respondents who were currently sexually active at the time of the interview.

In the economic empowerment dimension, the work status provides information on whether a woman was working at the time of the interview or in the 12 months prior to the interview, to account for the seasonality of the PSNP PW component. This includes every kind of work aside from housework (CSA 2001). Variations in the definition of work status will be exploited in the robustness checks. Moreover, I measure malnutrition based on the WFP definition, using the cutoff point for mild malnutrition as having a BMI lower than 18.5 (WFP 2005). Educational attainment is measured by the total years of received education, and to evaluate whether the PSNP alleviated women's time poverty, I look at the logarithmized time in minutes that a respondent needs to get to the next drinking water source and back.

[3]In 2016, the Ethiopian government launched an Urban Productive Safety Net Programme (UPSNP) as extension of the PSNP in 11 major urban cities of the country (Gebresilassie 2019)

To increase the precision of the estimates and reduce bias, the regressions include individual-level as well as district-level controls. Individual controls are taken from the DHS data sets and include a variable for electricity in the household, respondents' religion, five-year age cohorts and the sex of the household-head. District-level controls are collected from different data sources. Rainfall data is pulled from the CHIRPS data sets of the Climate Hazard Center and the US Geological Survey (CHC 2021). I compute the rainfall deviation from the long-term mean as percentage deviation of an annual 5-year mean from an annual 40-year mean per Woreda. The choice of 5-year averages aims at reflecting climatic shocks and thus possibly the economic development over the whole period preceding the respective survey round. This is reasonable, given that I observe many long-term outcomes that are unlikely to be affected by shocks in the same year and given that the gaps between the survey rounds are relatively large. Population density is taken from the NASA Socioeconomic Data and Applications Center and the CIESIN at the Columbia University, which provide UN WPP-adjusted population estimates that are based on the Ethiopian Census of 2007 (CIESIN 2017; CIESIN 2018). The rest of the Woreda-level controls, literacy rate, poverty rate and an infrastructure index, are based on DHS data. Details on the construction of outcome and control variables are provided in the appendix.

4.2 Timing of Surveys and PSNP Roll-out

The years of the DHS survey rounds that form my panel are 2000, 2005, 2011 and 2016. Ethiopia's Productive Safety Net Programme was planned to start in January 2005, whereas the fieldwork of the 2005 DHS survey has been carried out from April to August of 2005 (ICF 2021a). This overlap of the PSNP programme start and the 2005 DHS fieldwork prevents me from using the 2005 round as pre-treatment sample. Therefore, the 2000 survey round is used as pre-treatment sample throughout the main analysis, while the post-treatment sample is a pooled sample of the 2011 and 2016 DHS rounds. However, it remains to mention that the overlap of the 2005 survey and the PSNP start is small, particularly given the fact that I am mostly interested in long-term outcomes. Almost all variables that I observe can be expected to evolve only after a considerable amount of time, like the construction of better water infrastructure, social norms, or fertility (for which changes can be observed 9 months after the programme start at the earliest). Moreover, as mentioned before, the PSNP implementation finally started slightly delayed in April 2005, which reduces the overlap to only some months.
Therefore, I will include the 2005 survey in one of my robustness checks, to run the same analysis with the 2005 round as reference period.

5 Results

In this part I first present summary statistics for the outcome variables used in the analysis, as well as for the individual and Woreda-level control variables. Afterwards, the results of the main empirical specification are presented, and several robustness checks are carried out to check for the validity of the

estimates and to look for heterogeneous impacts. Finally, I discuss the obtained results in light of the PSNP programme design, the related literature and possible threats to validity.

5.1 Summary Statistics

At first, Table 1 presents the raw means of all women empowerment outcomes that are evaluated in this thesis, listed for the pre-treatment period in 2000. Indicators of women's social and economic empowerment are measured as described in the previous chapter. As one can see from the table, there are little noteworthy differences at baseline in the observed outcomes between Woredas that were targeted by the PSNP and those that were not targeted. Educational attainment, literacy and the share of women using contraceptives were lower in targeted Woredas at baseline, whereas more women approved wife beating (indicated by the "Violence" variable) and more women worked in the treatment group. However, none of these differences between the two groups is statistically significant. The only exception is the time that women needed to collect water. In PSNP Woredas, respondents needed significantly more time to get to the next drinking water place and back. This finding could reflect that the treatment group may have been poorer at baseline than control Woredas due to the programme selection criteria, and thus also worse equipped with infrastructure.

Table 1: Pre-treatment Means of Outcomes by Sub-sample

	(1) Control	(2) PSNP	(3) Difference	(4) P-Value	(5) Obs. Control	(6) Obs. PSNP
Fertility	0.423	0.436	-0.013	0.635	1859	2563
Contraception	0.066	0.057	0.008	0.680	987	1202
Violence	0.490	0.562	-0.073	0.248	1775	2439
Marriage	0.805	0.787	0.018	0.507	1700	2270
Malnutrition	0.391	0.384	0.008	0.814	1599	2187
Literacy	0.134	0.113	0.021	0.475	1823	2556
Work	0.590	0.704	-0.114	0.154	1859	2563
Water Time	3.101	3.797	-0.696***	0.001	1858	2563
Education	0.591	0.437	0.154	0.301	1859	2563

Notes: This table provides the baseline means of the variables evaluated in the paper for treatment and control group, as well as clustered t-tests. Means are computed from the pre-treatment sample, which is based on the DHS survey round of 2000. The sample is restricted to the rural population and to women aged 15-49. Standard errors are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

In general, the results shown in Table 1 point to a low level of women empowerment in rural Ethiopia at baseline, as measured by the variables here. Striking are for instance the low share of contraceptive users among sexually active women (6.6% and 5.7%, respectively), the high share of women who indicated that wife beating is justified when a wife refuses to have sex with her husband (49% and 56.2%), and a high share of malnourished women. In both groups, more than one third of the respondents were below

the cutoff for mild chronic energy deficiency. It is also important to note that education is not a binary variable here, but is measured in the total years of education received. Thus, respondents received on average only about half a year of education, which is manifested by the low level of literacy (13,4% and 11,3%). However, the share of women who were working at the time of the interview or in the preceding 12 months is high, particularly in the treatment group, where about 70% of women engaged in any kind of work aside from housework.

Figure 3 shows the raw means of the above-mentioned variables over the whole period of the four DHS survey rounds, plotting the trends from 2000 to 2016. Appealing in this respect is the steep and steady increase in many important outcomes in both groups, such as education, literacy and the use of contraceptive methods. Thereby, the means suggest that literacy eventually improved stronger in the treatment group and malnutrition decreased stronger in the control group, for instance. Importantly, Figure 3 also suggests mainly similar or parallel trends in the years between 2000 and 2005, except for women's attitude towards sexual violence. These graphs only show the raw means and the PSNP started a few months before the 2005 DHS survey, so this finding cannot be interpreted as absence of different pre-trends. However, given that the overlap of PSNP start and 2005 survey is small, it may be reassuring to see that almost all outcomes have not followed completely different paths in their raw means in the period between 2000 and 2005, and were thus also likely to have followed similar trends in the years that preceded the programme start.

Similar to the previous table, the baseline means for all individual and Woreda-level controls that are used in the analysis are presented in Table 4. As mentioned before, this includes indicators for poverty, infrastructure, literacy, population density and rainfall at Woreda-level, as well as religion, sex of the household head, electricity-access or age-groups at the individual-level. Significant differences in pre-treatment means of Woreda-level controls can be reported for rainfall, population density and the infrastructure index. Annual precipitation is generally much higher in control Woredas, as reflected in both the 2000 figure of annual rainfall and the 40-year average. Likewise, the five-year percentage-deviation in annual rainfall from the long-term mean was significantly different between treatment and control Woredas in 2000. Moreover, targeted Woredas were significantly poorer equipped with infrastructure at baseline, measured by the access to drinking water and electricity. Findings for both rainfall and infrastructure seem to be in line with the programme selection criteria of targeting particularly food insecure Woredas, given that drier Woredas are more likely to be agriculturally less productive and more sensitive to shocks. The low infrastructure level in treatment Woredas could point to a higher level of poverty. Also, the fact that the 5-year rainfall deviation was significantly different between both groups in 2000 suggests that rainfall shocks may indeed have affected treatment and control groups differentially, also over a longer period of several years. Moreover, targeted Woredas are on average significantly more densely populated than Woredas of the control group, while the baseline literacy rate tends to be higher in the control group and the poverty rate higher in targeted Woredas. However, differences in literacy and poverty rate are insignificant.

A more precise image of rainfall trends and climate hazards is provided by Figure 2. It presents the time trends of annual rainfall in PSNP and control Woredas back to 20 years before the last DHS survey round and illustrates some of the findings from Table 4. Annual rainfall is much higher in the control group than in the treatment group, and rainfall deviations from the long-term mean are partly varying between the two groups. As explained before, Ethiopia was repeatedly hit by droughts in the last decades. For instance, in 2015 the country was hit by the strongest drought of the past 50 years, which caused hunger among at least 10 Mio. Ethiopians (Alem Mera 2018). Interestingly, Figure 2 shows that targeted Woredas were much more affected by this drought, which occurred just before the last DHS survey round. This major deviation from the rainfall mean could have affected at least some of the women empowerment outcomes evaluated in this study, particularly nutrition or women's work status. Alem Mera (2018) points out that droughts in Ethiopia affected nearly all sectors and its social and economic impacts include migration, water conflicts and declines in GDP. The rainfall deviation thus seems to be an important control variable, also to serve as exogenous measure for the general economic development in rural Ethiopia.

All in all, the summary statistics point to the importance of controlling for the listed sets of variables, in order to capture possibly differential trends between treatment and control group. These variables can be expected to be correlated with the evaluated outcomes of women empowerment. At the same time, we could see here that these variables are partly differing a lot between treatment and control group, which could potentially bias the Diff-in-Diff results when they are not included in the analysis.

5.2 Impact of the PSNP on Women's Social and Economic Empowerment

Next, I present the main results of the Diff-in-Diff and triple differences specification as described in chapter 3, separately for the social and economic dimension of women empowerment. All binary outcome variables are thereby estimated in a LPM model. For the main specification, the DHS 2000 round serves as pre-treatment sample and the post-treatment sample is a pooled sample of the DHS 2011 and 2016 surveys. Furthermore, results are additionally presented for the single post-treatment years in Figure 4. This allows to observe whether the results are driven by one survey round in particular. Also, it is possible that the slight change in treatment assignment after 2009 that was discussed in chapter 3 is rather an issue for the 2016 than for the 2011 round.

Table 2 presents the effects of the Ethiopian Productive Safety Net Programme on social empowerment of women, here measured with a focus on women's control over their body. Columns 1 to 3 show the Diff-in-Diff coefficients for the specification that includes time fixed effects and Woreda fixed effects only, columns 4 and 5 present the respective Diff-in-Diff-in-Diff coefficients for early childbearing and early marriage. Columns 6 to 10 present OLS regression results for the same outcomes, now including the set of time-invariant and time-varying Woreda-level controls, as well as the set of individual-level controls.

Table 2: Main Results: Social Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Fertility	(2) Contra-ception	(3) Violence	(4) Early Childbearing	(5) Early Marriage	(6) Fertility	(7) Contra-ception	(8) Violence	(9) Early Childbearing	(10) Early Marriage
PSNP × Post	-0.0226	-0.0411	0.0316			0.0386	-0.147**	0.129		
	(0.0353)	(0.0441)	(0.0770)			(0.0419)	(0.0602)	(0.0890)		
PSNP × Post × Young				-0.0491	-0.0874				-0.0830	-0.0722
				(0.0567)	(0.0626)				(0.0578)	(0.0626)
Observations	12,633	6,414	12,246	12,633	11,411	12,584	6,387	12,200	12,584	11,370
R^2	0.020	0.141	0.062	0.065	0.353	0.229	0.181	0.076	0.232	0.481
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's social empowerment. Individual controls are age-group, electricity, sex of the household head, religion, Woreda-level controls include the rainfall-deviation and its square, as well as baseline values of the population density, literacy rate, poverty rate and an infrastructure index, interacted with a post-dummy. The analysis is based on the DHS survey rounds of 2000, 2011 and 2016, the pre-treatment sample is the DHS 2000 round, the post-dummy equals one for the years 2011 and 2016. Samples are restricted to women aged 15 to 49 and to the rural population. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

The overall results of Table 2 suggest that the PSNP had limited to no effects on the social empowerment of women. Although the coefficients of fertility, early childbearing and early marriage are negative, they are non-significant, as are the coefficients for contraception and violence perception in the base specification. Most of the outcomes are robust to adding the set of control variables and remain insignificant. The variable capturing women's use of contraceptives is even significantly negative once controls are included. Thus, the likelihood that women use any contraceptive method increased by 14.7pp less in the treatment group, compared to the control group. Figure 4 presents the results for the years of the pooled post-period separately. Apparently, the significant negative results for contraceptive use are driven by a sharp decline in 2011, which fades out in 2016. Moreover, the likelihood of women approving wife beating by the husband increased significantly in 2016, 11 years after the programme start. Interestingly, one can observe that the coefficient for early childbearing was significantly negative in 2011. The likelihood to have given birth in the three years preceding the survey was thus relatively lower among young women in treatment Woredas. All in all, although the coefficients lack statistical significance in most of the cases, it should be stated that the DDD coefficients for early childbearing and early marriage are negative throughout all specifications and all single survey rounds, and also show low p-values in the specification with all controls.

Table 3: Main Results: Economic Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Mal-nutrition	(2) Literacy	(3) Work	(4) Water Time	(5) Education	(6) Mal-nutrition	(7) Literacy	(8) Work	(9) Water Time	(10) Education
PSNP × Post	0.0762**	0.0330	-0.145**	-0.0968		0.171***	-0.0854	-0.0960	0.149	
	(0.0339)	(0.0341)	(0.0645)	(0.209)		(0.0470)	(0.0527)	(0.0792)	(0.143)	
PSNP × Post × Young					0.329					0.107
					(0.375)					(0.420)
Observations	10,738	12,371	12,630	12,613	12,633	10,692	12,326	12,581	12,568	12,584
R^2	0.045	0.105	0.188	0.204	0.245	0.066	0.247	0.214	0.263	0.348
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's economic empowerment. See notes of Table 2 for more details. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Estimation results are presented in the same way for the economic empowerment dimension in Table 3. Results for malnutrition and women's time poverty are also reported in this table. Firstly, no significant findings can be stated for literacy, education and the time to collect water, which also remain insignificant when I add the set of control variables. Although non-significant, the coefficient for literacy becomes even negative once controls enter the specification. Apparently, the PSNP was neither effective in improving women's education, nor in reducing the time women need to collect drinking water through an improvement in infrastructure. A surprising finding at this point is that malnutrition among women appears to have increased significantly in targeted Woredas relative to the control group. This finding is even stronger when I add the control variables which include rainfall shocks and can thus account for the incidence of droughts. Another interesting result can be reported from women's work status in the base specification, which suggests that the likelihood to work or have worked in the past year decreased significantly in PSNP Woredas, relative to the control group. However, although the coefficient remains negative at approximately 10pp, statistical significance does not survive in the specification that includes control variables. Figure 4 shows that the significant results for malnutrition remain constant across the two post-treatment periods. Also, there was a significant negative decline in the work status in the long run, whereas there was almost no change between 2000 and 2011. Since it is unlikely that the PSNP affected women's work status only after a considerable amount of time, this finding suggests that results pointing to a decline in women's likelihood to have worked should be interpreted with caution.

The previous analysis has presented the main results for the Diff-in-Diff and triple differences estimations, using a rich set of control variables and presenting results for both a pooled post-period and single years. Whereas the first general picture suggests that the PSNP did not improve most of the used indicators for women's economic and social empowerment, there is tentative evidence that the programme could have served as leverage to reduce the incidence of early childbearing. The partly intriguing results

that have been found in this section will be subject to further investigation in the subsequent part.

5.3 Robustness Checks

5.3.1 Event-study using 2005 as Reference Period

The most important identifying assumption in a Diff-in-Diff approach is the common trends assumption, which presumes that treatment and control group would have followed the same trends in outcomes in the absence of the policy (Angrist and Pischke 2008). In this thesis, the common trends assumption is addressed by including different Woreda-level controls that shall capture differential trends between the two groups, as well as a triple differences approach for some outcomes, which can be expected to further address some group-specific trends. An important additional step in providing credibility to common trends would be a Placebo-test, using two samples that precede the timing of the treatment, for instance. In the setting at hand this is a difficult undertaking, given that only the 2000 survey was fielded before the implementation of the Productive Safety Net Programme. As described in Chapter 4, the 2005 DHS survey slightly overlapped with the start of the PSNP programme and was fielded some months after its start. Whereas this timing makes the sample inappropriate for both using it as post-period or using it as pre-treatment period in a clean Placebo-test, it is nevertheless of interest to run the analysis again, this time including the 2005 survey round and using it as reference period. This allows to observe trends before and after the year in which the PSNP was implemented.
Thereby, it is important to keep in mind that many outcomes that are evaluated in this analysis are long-term outcomes which are unlikely to be affected by a policy within several weeks or months, such as fertility, improvements in infrastructure, education or literacy. In this respect, observing trends from 2000 to 2005 could give a first impression of whether long-term outcomes were likely to follow different paths between treatment and control group in the years before the PSNP started. Also, using the 2000 sample as pre-treatment period with several years between sample fieldwork and the programme start can leave room for imprecisions due to changes in the years before treatment, which makes it additionally important to carry out this robustness check. Table 5 and Table 6 show the results of this specification, which uses 2005 as reference period. Again, the coefficients of all single years are shown in Figure 5, now including the 2005 sample.

First of all, the graphs and regression tables show that there were mostly no significantly different changes in outcomes between targeted and non-targeted Woredas in the time between 2000 and 2005. Only malnutrition, early marriage and contraceptive use show significant coefficients in the period from 2000 to 2005 once all controls are included, at a low significance level. Although a clear conclusion with respect to the common trend assumption cannot be drawn, the findings imply that differential trends in outcomes before the programme start were not an issue at least for the long-term outcomes like fertility, education, literacy or the time to collect water.
However, the graphs in Figure 5 and the corresponding tables can still shed more light on the findings from the previous section, comparing the post-treatment years to 2005, briefly after the programme start.

Importantly, the graph for early childbearing casts doubts on the partly significant finding of a decrease in early childbearing through the PSNP. Given that this variable cannot be affected by the programme within a few months, the non-significant and smaller decrease from 2005 on weakens the argument for a significant impact on early childbearing. Apart from that, the results in the social empowerment dimension remain mostly similar compared to before. However, the significant decline in contraceptive use disappears when I use the 2005 round as reference period and the sexual violence variable is now significantly positive for the pooled post-period, which is mainly caused by a sharp increase in 2016.
Likewise, in the economic empowerment dimension the non-significant results found in the previous chapter for work status, education, the time to collect water and literacy are robust to using 2005 as reference period. It can be mentioned, however, that the graph for education may suggest an increase in educational attainment in 2011, although this finding lacks statistical significance. Interestingly, women's likelihood of being malnourished now shows only a weak significant relative increase in 2011 in PSNP-Woredas, which fades out in 2016. Apparently, the results from before are mainly driven by the relative increase in malnutrition from 2000 to 2005 in targeted Woredas. In either case this finding does not change the fact that the PSNP appears to have been unsuccessful in mitigating women's malnutrition.

To sum up, the results in this section should be interpreted with caution, taking into account the particular timing of the 2005 DHS survey and the PSNP programme start. Keeping in mind that most variables are long-term outcomes, some of the findings can suggest that a violation of the common trends assumption was likely not to be an issue for many outcomes. On the other hand, some of the findings from before, such as for early childbearing or malnutrition, could be at least partly driven by differential trends that started to evolve before the treatment.

5.3.2 Investigating Treatment Intensity

As mentioned in chapter 2, while the Productive Safety Net Programme was rolled-out at Woreda-level, not everyone within a Woreda was eligible to participate. Instead, Kebeles and households were subject to a further targeting process once a Woreda was assigned to the PSNP, aiming to select the most food insecure households (World Bank 2012). Moreover, the scale of programme coverage was not homogeneous across targeted Woredas, with differences in the intensity of treatment (Gazeaud and Stéphane 2020). One may thus worry that the obtained results could be biased due to a low level of PSNP coverage in some Woredas that are assigned to the treatment group, but were barely covered by the programme. Therefore, I exploit differences in treatment intensities in PSNP Woredas, drawing on data provided by Gazeaud and Stéphane (2020) (based on FAO (2007)). It should be noted that the data for treatment intensity is only available for 2006 and that some Woredas that joined later were not assigned an intensity. In this case, Gazeaud/Stéphane assigned the average intensity level. Also, they provide findings which suggest that the intensity levels have remained stable over time (Gazeaud and Stéphane 2020). Treatment intensity is defined as the percentage of the population that is targeted by the PSNP within a Woreda and is grouped into broad categories. Figure 1 shows that the intensity levels range from a low 2%-13% in some

of the southern Woredas to 43%-65% in the north-eastern districts in my sample. Following Gazeaud and Stéphane (2020), I use the data on treatment intensity in two ways. First, I replace the treatment dummy by an ordered categorical variable of intensity levels, which I treat as continuous. Thus, the new treatment variable takes values from 0 to 4, where 0 is assigned to Woredas of the control group. The results now show the effect of an increase in treatment coverage by one level, instead of the direct comparison between targeted and non-targeted Woredas. Secondly, using a treatment dummy again, I run the main specification of section 5.2 with a sample that includes only the most intensely targeted Woredas (keeping intensity levels 3 and 4). The control group is then compared to all Woredas in which 26%-42% or 43%-65% of the population was targeted by the PSNP. This reduces the size of the treatment group to 29 Woredas.

Table 7 to Table 10 show that the results obtained in chapter 5.2 are mostly robust to exploiting the different scales of PSNP roll-out across Woredas. Using the continuous treatment variable, I again find no significant effects on fertility, attitude towards violence and early marriage, whereby the coefficient of the violence variable now has a negative sign. As before, early childbearing shows a negative but insignificant coefficient with the pooled sample. In line with previous findings, the use of contraceptives decreased in more intensely treated Woredas relative to less treated districts, now even significant at the 1%-level. Very similar results in terms of signs and significance for the social empowerment dimension are obtained using the sample of most treated Woredas, shown in Table 8. Here, early marriage shows a significantly negative coefficient in the base specification without controls. All in all, the triple differences coefficients that measure the effect on early childbearing and early marriage are large and negative throughout all specifications again. The robustness of the model to variations in the definition of the treatment variable and the sample also applies to the economic empowerment dimension (Table 9 and Table 10). Apart from a significantly (10%-level) positive increase in the time to collect water with the continuous treatment variable and controls, coefficients for literacy, education, water time and work status remain insignificant for both using a continuous treatment variable and cutting the sample to most treated districts. The findings of a relative increase in malnutrition in the treatment group can be confirmed as well.

5.3.3 Probit Regression Model

Most of the outcomes that are evaluated in this paper are binary variables, where coefficients can be interpreted as the change in the probability that the dependent variable takes the value one, and which are estimated in a Linear Probability Model. To shed more light on the robustness of the results of the previous sections, I investigate effects of the PSNP on the binary outcomes in the analysis using a Probit Model and Maximum Likelihood Estimation techniques instead of a LPM. As Table 11 and Table 12 show, the probit model yields similar to same results as under the LPM. As before, no significant effects can be stated for fertility, violence-perception, early marriage, early childbearing and women's work status in the final specification. In the base specification, early marriage now has a significantly negative coefficient, but it does not hold for the specification where I add the control variables. The relative decrease in the use of contraceptive methods remains robust, as well as the relative increase in malnutrition. One exception

is that the probit model suggests that literacy increased significantly less in targeted Woredas compared to the control group.

5.3.4 Impact Heterogeneity: Work status and age cohorts

The previous sections revealed many negative coefficients for women's work status which were even significant in some specifications. Given that I evaluate a Public Employment Programme, this is a surprising result. It is certainly possible that additional income through the safety net could have caused negative externalities on women's employment, reducing the need to engage as additional workforce of the household due to an improved income situation. However, in this section I want to take a closer look at the definition of women's work status. Although the measure of employment used in the main analysis excludes housework, the definition does not rule out the possibility to have engaged in unpaid work or to have worked for family members.

Therefore, I evaluate the impact of the PSNP on four refined definitions of women's work status, as documented in Table 13. Firstly, the same work variable is used as before, now excluding women younger than 18 from the analysis (Column 1). This is the official age threshold that qualifies for participating in the PSNP PW component, and I thus want to rule out that the officially non-eligible youngest women biased the results. Secondly, I use a work status variable that excludes unpaid work from the definition, such that the dummy equals one only when a respondent worked and was paid either in cash, in kind or both for her work (Column 2). Thirdly, I directly examine the likelihood to have engaged in paid work, compared to unpaid work, which only concerns working women (Column 3). Finally, I examine the likelihood to have worked for others, compared to have worked for either family members or being self-employed (Column 4).

In a second part of this robustness check, I am interested in comparing PSNP effects on the work status between younger and older women. In the main specification, there was some significant result for a negative effect on early childbearing, and coefficients for both early marriage and early childbearing were negative throughout all specifications. At the same time, there were no results for an increased likelihood to work, which I assumed to serve as possible leverage to delay childbearing or marriage in my conceptual framework. However, it is possible that the PSNP PW particularly hired younger women and thus affected early childbearing and early marriage, which I will evaluate with a triple differences estimation again, comparing the work status of a young group to the rest. In contrast to before, I define the young cohort as women between 18 and 22 or between 18 and 24, since women under 18 were officially not allowed to participate in the Public Works component.

Table 13 presents the results for the simple Diff-in-Diff and shows that the before obtained results are robust to excluding women under 18 from the analysis (Column 1). A similar (insignificant and negative) result is obtained for the work status, when unpaid work is excluded from the definition. Although the coefficient is large and positive, there is no significant finding that the PSNP increased the likelihood to be paid for one's work. However, apparently the PSNP increased the likelihood to have worked for

others instead of being self-employed or working for relatives. This finding is significant for both the base specification and the estimation with all control variables. The results imply that, whereas the programme did not increase the overall likelihood for women to work, it could have crowded out the work for family members or self-employment. Further, this could imply that the programme particularly targeted women who were already engaged in some kind of paid or unpaid work before the programme start. Interestingly, Table 14 (here only displaying the specification with control variables) provides evidence that there seem to be differential impacts of the PSNP on the work status of younger women. Including the set of control variables, women from 18 to 22 in PSNP Woredas were significantly more likely to work, relative to the other age groups and to control Woredas. This finding is even stronger when unpaid work is excluded from the definition of the work status. Furthermore, effects are stronger when I vary the definition of the young cohort and compare women aged 18 to 24 to the rest. Young women of the treatment group who are in a PW-eligible age are more likely to work, which could suggest that the Safety Net Programme particularly hired younger women to participate in the Public Works component.

Finally, particularly given these results, it is of interest to check for impact heterogeneity in early childbearing and marriage across different definitions of the young cohort. To measure early childbearing and early marriage, I defined the young group as women between 15 and 22. However, the results that revealed many large negative coefficients with low p-values, might be sensitive to changes in the age group. I will assess impacts on early childbearing and early marriage for three more definitions of the young cohort and show the results in Table 15, now comparing effects for women between 15 and 20 (columns 1-2), women between 15 and 24 (columns 3-4) and women between 18 and 22 (columns 5-6) to the rest. Given that I expect possible impacts on early childbearing and marriage to be particularly driven by young women's work opportunities, but only women aged at least 18 were officially eligible for the PSNP Public Works component, the last specification is of special interest. It can show whether the significant relative increase for the youngest group of PW-eligible women (18-22) in the likelihood to work is accompanied by a significant relative decrease in fertility and marriage for this age group.
As we can see from Table 15, there is neither evidence for such a significant effect, nor for impact heterogeneity across different definitions of the age cohorts in general. The coefficients remain negative for all variations in age groups. However, although they are relatively large and have low p-values for early childbearing, the results remain insignificant. Thereby, the largest negative coefficient of fertility can be found for the group of 18 to 22 year old women.

5.4 Discussion

The previous sections presented the results of the main Diff-in-Diff and triple differences specification for a range of outcomes related to the social and economic empowerment of women, as well as several robustness checks. Ethiopia's Productive Safety Net Programme (PSNP) appears to have had limited impacts on the empowerment of women in a low-income setting, as measured by these outcomes. Coefficients for early childbearing and early marriage are negative throughout all specifications and there is some evidence

that early childbearing responded significantly negatively to the PSNP, but in most cases coefficients for both outcomes remain insignificant. Moreover, malnutrition increased significantly in treatment Woredas relative to the control group, and the use of contraceptive methods increased significantly less in the treatment group in most specifications. The obtained results are mostly robust to several adjustments, such as exploiting the programme's treatment intensity, using a Probit model for the binary variables or including the 2005 sample as reference year. However, some robustness checks suggest that the programme could have shifted work from self-employment to working for others and that it increased the relative likelihood to work among the youngest eligible women.

I interpret these limited impacts in bringing about changes in women empowerment in two ways. First, my findings could point to the fact that leverages of the programme to empower women did not work as I suggested in the conceptual framework. For instance, more income security through the safety net did not appear to make parents invest more in their daughter's education. Also, as many of my assumptions in chapter 2 rely on the idea that social empowerment could have been spurred by an increase in female work, the results would call the importance of female employment into question in this respect. However, my study shows that the overall work status of women between 15 and 49 did not positively respond to the PSNP implementation. This can point to the fact that the PSNP was not successful in integrating women into its Public Works component, or that the PSNP even caused negative employment externalities by reducing the need of additional workforce in the household through increased income. Or, the overall work status of women did not change significantly since recruited women already engaged in some other kind of work before. Robustness checks that refine the work definition suggest that rather the latter applies, providing significant results for a shift from self-employment or working for family members to working for others. Also, first qualitative assessments reported positive experiences in terms of women's participation and integration into the programme (World Bank 2012).
As a consequence of the non-increase in women's overall likelihood to work, social empowerment outcomes cannot be affected through the channel of female employment, which could partly explain a lack in significant impacts on fertility or women's attitude towards violence.
Secondly, many of the findings in this study could imply that the PSNP was not implemented efficiently enough to make a difference, particularly in terms of gender sensitivity. Field evidence partly documented difficulties in gender-responsive programme implementation. Whereas no problems were reported in integrating women into the programme in general, first assessments found that in some places the programme's timing coincided with peak agricultural seasons and that extended physical labour was sometimes conflicting with other productive activities and women's domestic responsibilities (Sharp et al. 2006; Holmes and Jones 2011).

In general, the findings of my thesis add to and corroborate some of the prior empirical work on the impacts of the PSNP in Ethiopia. In Tigray, Gelegay and Lecoutere (2019) find almost no positive and some negative effects of the PSNP on women's decision-making power about economic aspects in the household,

and null results for effects on an aggregated empowerment score. Only the effect on women's access to credit is significantly positive. Gilligan et al. (2009) found no lasting impacts on a range of outcomes, which include food security, when treatment was defined as having received any payment for participating in the PSNP Public Works (without transfers of the OFSP). Furthermore, Gazeaud and Stéphane (2020) report null results for effects of the PSNP on agricultural productivity. My findings for the unchanged overall work status of women aged 15–49 could also corroborate the lack of significant impacts on overall employment through the PSNP that Schinaia (2016) found in this context. On the other hand, Imbert and Papp (2015) find in another setting that the Indian NREGA decreased private sector work, including self-employment and domestic work. My finding for a shift from self-employment and working for family members to working for others could be regarded in light of their finding.

The negative effects on nutrition and contraception, however, remain surprising findings. As shown in the robustness checks, these results should be interpreted with caution, given that differential long-term trends between treatment and control Woredas cannot be entirely ruled out. However, the result in nutrition may be explained to some extent by an increased exposure of women to physical strain and less time through the Public Works. Qualitative evidence documented frequent complaints that the PSNP did not adapt the workload accordingly. Some interviewed female participants reported difficulties to manage childcare and domestic responsibilities alongside public work and complained about the heavy workload (Sharp et al. 2006). Increased exposure to physical strain and less time for child- and self-care might partly explain the increased incidence of chronic energy deficiency relative to the control group.

Although many of the evaluated outcomes seem to be unaffected or even negatively affected by the PSNP, it remains to mention that the coefficients of the triple differences for early childbearing and early marriage were negative throughout all specifications and robustness checks, showing many low p-values. Some evidence was found for a significant decrease in early childbearing in 2011. Robustness checks that used the 2005 sample as reference period suggested that these results should be interpreted with caution. However, given that I also found significant increases in female work among the youngest eligible women, there can be expectations that the PSNP delays or suspends early childbearing and marriage, possibly by disproportionally hiring younger women for Public Works. In this case, the PSNP can serve as leverage to increase young women's future employment and education opportunities. As for instance Hoddinott and Mekasha (2020) find some evidence for reduced overall fertility in such a context, and this thesis could not bring up clear significant results in this respect, it could be an important avenue for future research to further investigate impacts of the PSNP and other Public Employment Programmes on early childbearing and early marriage.

Finally, the discussion of my analysis must also bring up possible threats to validity and limitations of the data. First, it should be mentioned that it is crucial for the common trend assumption that no similar programmes in other states or Woredas were running parallel to the PSNP during the evaluation period. It cannot be entirely ruled out that states such as the Eastern regions Benishangul Gumuz and Gambella, which are not targeted by the PSNP, received other support. However, it is more likely that such policies

took the form of food aid rather than large-scale employment programmes, which should then not affect most of the evaluated women empowerment outcomes. A possible bias could also arise if the PSNP causes migration flows when poor households from neighbouring but non-targeted Woredas are attracted by the program. In the context of migration, Hoddinott and Mekasha (2020) present evidence that the PSNP decreased out-migration of young women and Gazeaud and Stéphane (2020) provide suggestive evidence that the PSNP did not significantly affect immigration to beneficiary Woredas. Therefore, inter-rural migration is likely not to be an issue in this setting.

As discussed in chapter 3, an increase in programme coverage after 2009 could pose a threat to validity when Woredas that were only targeted after 2009 are wrongly assigned to the control group, since the targeting data is of 2009. However, programme documents report only a slight increase in Woredas after 2009 (319 in the time between 2010 and 2015, compared to 291 in 2009) (World Bank 2016). It is therefore unlikely that some of these added Woredas are represented in the 32 Woredas of my control group.

Finally, particularities with respect to the timing of the surveys and the data should be considered in the interpretation of my results. Whereas the 2005 survey does not represent an adequate pre-programme sample, the 2000 survey was fielded several years before the programme implementation. Thus, the possibility that some of the observed effects are partly driven by developments between 2000 and 2005 cannot be entirely ruled out. Also, one must recall that, whereas impacts are evaluated at Woreda-level, not everyone within a Woreda finally participated in the PSNP. This could leave room for slight imprecisions, or for underestimations of the programme impact to a certain extent. In terms of the data situation it should also be mentioned that the number of available individual observations and Woredas has been greatly reduced in order to construct a balanced Woreda-year panel, given that the DHS program did not sample the same Woredas in each survey round.

To conclude, the mentioned possible threats to validity and data limitations do not appear to be a major issue for the identification strategy of this study. However, it would be interesting to run this analysis or a similar specification using other data sources and other timings. For instance, it could be important to exploit the fact that the PSNP was rolled-out in several steps, drawing on yearly data for the treatment assignment and accounting for variations in treatment over time and the duration of programme exposure[4]. This was not in the scope of this paper, but could be an interesting avenue for future research. Another way of approaching this topic could be to put a stronger focus on the agency dimension of women empowerment in evaluations of Public Employment Programmes, which was not possible with the data at hand. For instance, Brauw et al. (2014) found significant improvements in women's intra-household decision-making power for the Brazilian CCT Program Bolsa Familia in urban areas. As the Ethiopian Government recently launched an Urban Safety Net Programme (UPSNP) (Gebresilassie 2019), it could also be of interest to extend my analysis to urban areas. Finally, the external validity of the results obtained in this study for the PSNP should be investigated. The Ethiopian context is very specific and might not be easily compared to other settings like South Africa or India. To the best of my knowledge, impact

[4]This idea was brought up by an anonymous referee mentioned by Jules Gazeaud

evaluations of large-scale Public Employment Programmes in terms of women empowerment remain scarce, even for the MGNREGA in India. It is therefore of interest to apply studies like this to other contexts and evaluate which programme designs may be more effective in strengthening women's role in low- and middle-income settings.

6 Conclusion

Public Employment Programmes are a policy tool that gained popularity and is frequently used to support the working age poor and promote development, globally employing almost 70 million people, particularly in South Asia, South East Asia and Sub-Saharan Africa (McCord and Paul 2019). This thesis is one of the first papers to empirically evaluate impacts of a Public Employment Programme on the empowerment of women, focusing on one of the largest Social Protection Programmes in Africa, the Ethiopian PSNP. My study also adds in an important way to existing evidence of the efficiency and general impacts of this large-scale policy.

In the conceptual framework, I highlighted several mechanisms through which an employment and social protection programme like the PSNP could affect outcomes related to the empowerment of women. While women empowerment is generally a highly diverse concept, I put a particular focus on women's social empowerment, added by measures of education, work, nutrition and time poverty. Most importantly, I emphasized the possibility that the programme could serve as leverage to reduce fertility, early childbearing and early marriage, when opportunity costs of having children become higher through programme participation. Easing a household's budget, the programme could also affect investment decisions in girls' education.

Employing a Diff-in-Diff approach and using four cross-sections of the Demographic and Health surveys (DHS), this study finds limited impacts of the PSNP on the empowerment of women. Whereas most of the evaluated outcomes appear to be not significantly or even negatively affected by the programme, coefficients of early childbearing and early marriage are negative across all specifications, and in few cases significant. Moreover, while the PSNP did not increase the overall likelihood of women to engage in any kind of work, it seems that the programme crowded out self-employment and working for family members. Finally, the PSNP appears to positively affect the likelihood to work for the youngest eligible cohort of women.

The first general picture suggests that the Ethiopian Productive Safety Net Programme does not appear to be a game changer in promoting women empowerment in a low-income setting, as measured by the outcomes at hand. Although some of the results should be interpreted with caution, the findings can suggest that either the mechanisms of the programme did not affect women empowerment outcomes as expected, or the PSNP lacked an efficient and gender-equitable implementation. While empowering women was not the primary goal of the programme, this could still point to the need to implement large-scale programmes like this more efficiently in terms of gender-sensitivity. However, the negative coefficients for

early childbearing and marriage, combined with the increased relative likelihood to work for the youngest eligible women, could imply a potential role for the PSNP of serving as leverage to delay or suspend early childbearing and marriage. In this case, the employment programme could enhance young women's human capital trajectories and future labor market opportunities in an important way. Although lacking statistical significance in most cases, particularly this finding should motivate further investigation in the Ethiopian context or other settings. The discussed mechanism could not only empower women and make them less dependent on men, but could also serve as general channel to foster development and growth by promoting female employment and reducing fertility.

As large-scale Public Employment Programmes become of increasing importance and empowering women remains one of the central development goals, this topic offers rich possibilities for future research. Some of them have been suggested in the previous chapter and include a stronger focus on women's agency position or an extension of the analysis to urban areas. While the Ethiopian context still offers the possibility to exploit this topic further, it seems of particular importance to investigate the research question in other settings as well, as there is a growing number of similar programmes such as in Tanzania, Malawi or Mozambique (McCord and Paul 2019).

References

Alem Mera, G. (2018). "Drought and its impacts in Ethiopia". In: *Weather and Climate Extremes* 22, pp. 24–35.

Amaral, S., S. Bandyopadhyay, and R. Sensarma (2015). "Employment Programmes for the Poor and Female Empowerment: The Effect of NREGS on Gender-based violence in India". In: *Journal of Interdisciplinary Economics* 27 (2), pp. 199–218.

Angrist, J.D. and J. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford: Princeton University Press.

Anthony, M., M. Budde, C. Funk, G.J. Husak, M. Landsfeld, and S. Shukla (2021). "Enhancing the Application of Earth Observations for Improved Environmental Decision-Making Using the Early Warning eXplorer (EWX)". In: *frontiers in climate* 2, pp. 1–16.

Antonopoulos, R. (2007). "The Right to a Job, the Right Types of Projects: Employment Guarantee Policies from a Gender Perspective". In: *Levy Economics Institute Working Paper*. No. 516.

Baird, S., E. Chirwa, J. de Hoop, and B. Özler (2016). "Girl Power: Cash Transfers and Adolescent Welfare. Evidence from a Cluster-Randomized Experiment in Malawi". In: *" African Successes, Volume II: Human Capital"*. University of Chicago Press: pp. 139-164.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2020). "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa". In: *American Economic Journal: Applied Economics* 12 (1), pp. 210–259.

Banerjee, V.A. and E. Duflo (2011). *Poor Economics. A Radical Rethinking of the Way to Fight Poverty*. London: Penguin Books.

Berck, P. and S.B. Villas-Boas (2016). "A note on the triple difference in economic models". In: *Applied Economics Letters* 23 (4), pp. 239–242.

Berhane, G., D.O. Gilligan, J. Hoddinott, N. Kumar, and A. Seyoumg Taffesse (2014). "Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Programme". In: *Economic Development and Cultural Change* 63 (1), pp. 1–26.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How Much Should We Trust Differences-In-Differences Estimates?" In: *The Quarterly Journal of Economics* 119 (1), pp. 249–275.

Brauw, A. de, D.O. Gilligan, J. Hoddinott, and S. Roy (2014). "The Impact of Bolsa Família on Women's Decision-Making Power". In: *World Development* 59, pp. 487–504.

Burgert, C. R., J. Colston, T. Roy, and B. Zachary (2013). *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*. DHS Spatial Analysis Reports No. 7. Calverton, Maryland, USA: ICF International.

Center for International Earth Science Information Network (CIESIN), Columbia University (2017). Gridded Population of the World, Version 4(GPWv4): Administrative Unit Center Points with Population Estimates, Revision 10. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). URL: https://doi.org/10.7927/H46H4FCT [Accessed May 14, 2021].

— (2018). Documentation for the Gridded Population of the World, Version 4(GPWv4), Revision 11 Data Sets. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). URL: https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/documentation [Accessed May 14, 2021].

Central Statistical Authority [Ethiopia] and ORC Macro (2001). *Ethiopia Demographic and Health Survey 2000*. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Authority and ORC Macro.

— (2013). Standard Recode Manual For DHS 6. Funded by USAID. Addis, Ababa, Ethiopia and Rockville, Maryland: ICF [Distributor].

— (2000-2016). Demographic and Health Surveys (various) [Datasets]. Funded by USAID. Addis, Ababa, Ethiopia and Rockville, Maryland: ICF [Distributor].

Climate Hazard Center (2021). *CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations*. URL: https://earlywarning.usgs.gov/fews/climate-workshop [Accessed May 14, 2021].

Croft, T., A.M.J. Marshall, and C.K. Allen (2018). *Guide to DHS Statistics*. Rockville, Maryland, USA: ICF.

Devereux, S., R. Sabates-Wheeler, R. Slater, M. Tefera, T. Brown, and A. Teshome (2008). *Ethiopia's Productive Safety Net Programme (PSNP): 2008 Assessment Report*. URL: https://www.cashdividend.net/wp-content/uploads/2013/05/Devereux-et-al.-2008.pdf [Accessed May 14, 2021].

Duflo, E. (2012). "Women Empowerment and Economic Development". In: *Journal of Economic Literature* 50 (4), pp. 1051–1079.

Food and Agriculture Organisation (2007). *Percent of Population targeted for Productive Safety Net Program (PSNP) in 2006*. URL: http://www.fao.org/3/j9325e/j9325e00.htm [Accessed May 15, 2021].

Gazeaud, S. and V. Stéphane (2020). "Productive Workfare? Evidence from Ethiopia's Productive Safety Net Program". In: ffhalshs-03082420.

Gebresilassie, Y.H. (2019). *Ethiopia's Urban Productive Safety Net Programme: Its Impact on Households' Food Insecurity*. URL: https://www.uneca.org/sites/default/files/AEC/2020/presentations/ethiopias_urban_productive_safety_net_programme_-_its_impact_on_households_food_insecurity_.pdf [Accessed May 14, 2021].

Gelegay, M.D. and E. Lecoutere (2019). "Did conditional cash transfers in the Productive Safety Net Program empower women in Tigray, north-east Ethiopia?" In: *IOB Discussion Papers*. 2019.01.

Gilligan, D.O., J. Hoddinott, and A. Seyoum Taffesse (2009). "The Impact of Ethiopia's Productive Safety Net Programme and its Linkages". In: *The Journal of Development Studies* 45 (10), pp. 1818–1837.

Hoddinott, J. and T.J. Mekasha (2020). "Social Protection, Household Size, and Its Determinants: Evidence from Ethiopia". In: *The Journal of Development Studies* 56 (10), pp. 1818–1837.

Holmes, R. and N. Jones (2011). *Gender inequality, risk and vulnerability in the rural economy. Refocusing the public works agenda to take account of economic and social risks*. ESA Working Paper No. 11-13. London: Overseas Development Institute.

ICF (2021a). *Ethiopia Standard DHS, 2005. Survey Summary*. The DHS Program website. Funded by USAID. URL: https://www.dhsprogram.com/methodology/survey/survey-display-248.cfm [Accessed May 14, 2021].

— (2021b). *Using Datasets for Analysis*. The DHS Program website. Funded by USAID. URL: https://www.dhsprogram.com/data/Using-Datasets-for-Analysis.cfm [Accessed May 14, 2021].

Imbens, G. and J.M. Wooldridge (2007). *What's New in Econometrics? Difference-in-Differences Estimation. Lecture Notes 10*. NBER.

Imbert, C. and J. Papp (2015). "Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee". In: *American Economic Journal* 7 (2), pp. 233–263.

International Labour Organisation (2021). *Public Employment Programmes*. URL: https://www.ilo.org/global/topics/employment-intensive-investment/themes/pep/lang--en/index.htm [Accessed May 13, 2021].

Kabeer, N. (1999). "Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment". In: *Development and Change* 30 (3), pp. 435–464.

Lavers, T. (2019). "Social protection in an aspiring 'developmental state': The political drivers of Ethiopia's PSNP". In: *African Affairs* 118 (473), pp. 646–671.

Lowe, C., E. Ludi, M. Le Sève, and J. Tsui (2019). *Linking social protection and water security to empower women and girls*. Working paper 567. London: Overseas Development Institute.

McCord, A. and M.H. Paul (2019). *An Introduction to MGNREGA Innovations and their Potential for India-Africa Linkages on Public Employment Programming*. New Delhi: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) India.

Ministry of Agriculture Addis Abeba (2014). *Productive Safety Net Phase IV - Programme Implementation Manual*. Addis Ababa. URL: https://www.usaid.gov/sites/default/files/documents/1866/psnp_iv_programme_implementation_manual_14_dec_14.pdf [Accessed May 13, 2021].

Open Africa (2021). *Ethiopia Shapefiles*. URL: https://africaopendata.org/dataset/ethiopia-shapefiles [Accessed May 14, 2021].

Schinaia, G. (2016). *Labour market effects of social programmes: Evidence from Ethiopia's productive safety net*. Master's thesis, University of Oxford.

Sen, A. (1990). "More Than 100 Million Women Are Missing". In: *The New York Review of Books* 1990 Issue.

Sharp, K., T. Brown, and A. Teshome (2006). *Targeting Ehtiopia's Productive Safety Net Programme (PSNP)*. London: Overseas Development Institute.

Stecklov, G., P. Winters, J. Todd, and F. Regalia (2007). "Unintended effects of poverty programmes on childbearing in less developed countries: Experimental evidence from Latin America". In: *Population Studies* 61 (2), pp. 125–140.

Todd, J.E., P. Winters, and G. Stecklov (2010). "Evaluating the impact of conditional cash transfer programs on fertility: the case of the Red de Protección Social in Nicaragua". In: *Journal of Population Economics* 25, pp. 267–290.

UN Office for the Coordination of Humanitarian Affaires (2007). *Productive Safety Net Programme (PSNP) Woredas and Responsible Agency*. URL: https://reliefweb.int/map/ethiopia/ethiopia-productive-safety-net-program-psnp-woredas-and-responsible-agency-24-jul-2007 [Accessed May 14, 2021].

UN Women (2021). *Africa*. URL: https://www.unwomen.org/en/where-we-are/africa#:~:text=The%20African%20continent%20has%20demonstrated,Rights%20of%20Women%20in%20Africa [Accessed May 13, 2021].

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Fertility Report 2015 - Highlights*. (ST/ESA/SER.A/415).

Verbeek, M. (2008). "Pseudo Panels and Repeated Cross-Sections". In: *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Ed. by L. Mátyás and P. Sevestre. 3rd ed. Springer. Chap. 11, pp. 369–382.

World Bank (2012). *Ethiopia - Designing and Implementing a Rural Safety Net in a Low Income Setting: Lessons Learned from Ethiopia's Productive Safety Net Program 2005-2009*. Washington DC: World Bank.

World Bank. Social Protection Unit Africa Region (2010). *Implementation Completion and Results Report*. Report No: ICR00001676. URL: http://documents1.worldbank.org/curated/en/525401468024639387/pdf/ICR16760P098091e0only1910BOX358365B.pdf [Accessed May 14, 2021].

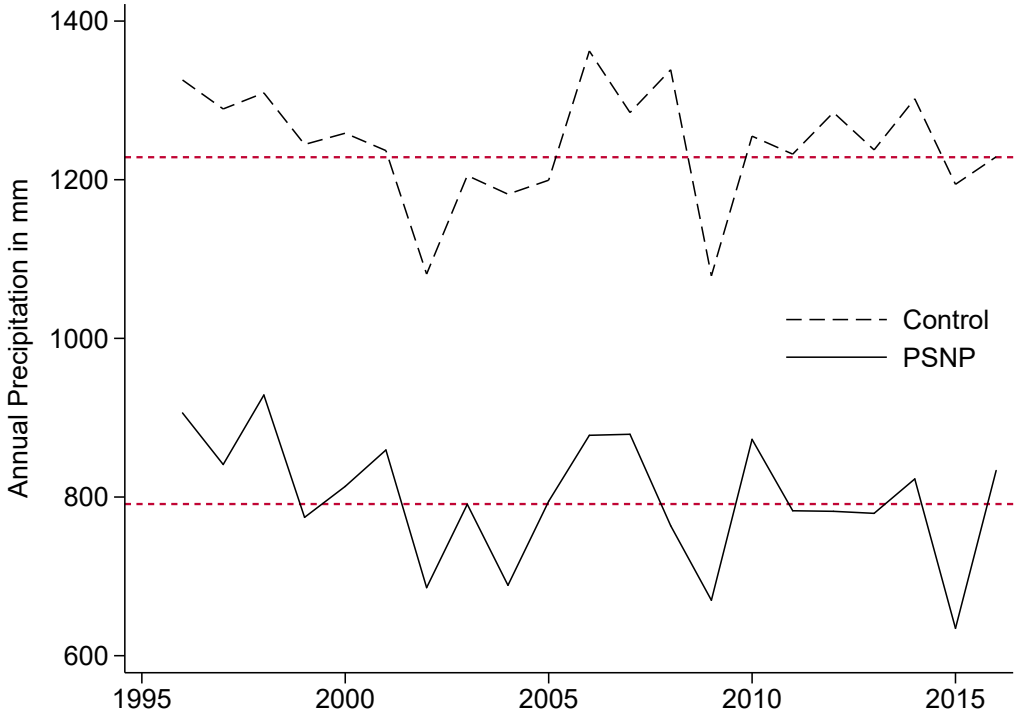
— (2016). *Implementation Completion and Results Report*. Report No: ICR00003659. URL: http://documents1.worldbank.org/curated/en/229421467954050889/pdf/ICR3659-P113220-Box396273B-PUBLIC-disclosed-7-6-16.pdf [Accessed May 14, 2021].

World Food Programme (2005). *A Manual: Measuring and Interpreting Malnutrition and Mortality*. Rome: World Food Programme.

Appendix

Tables and Figures

Figure 2: Annual precipitation from 1996–2016, in mm



Notes: This figure plots the annual rainfall over 20 years, separately for treatment and control group of the Woredas that constitute my panel. The upper red dashed line represents the 40-year mean in rainfall for the control group, the lower red dashed line represents the respective long-term mean in the treatment group. Rainfall is measured in mm.
Source: Author's elaboration based on CHC (2021) and World Bank (2010).

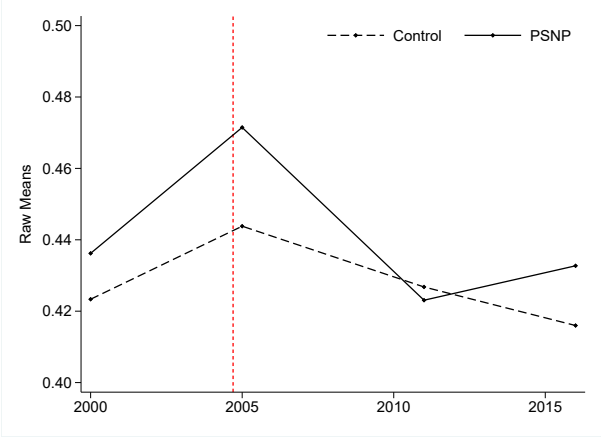
Table 4: Pre-treatment Means of Controls by Sub-sample

	(1) Control	(2) PSNP	(3) Difference	(4) P-Value
A. Woreda-level Controls:				
Poverty Rate	0.336	0.392	-0.056	0.460
Literacy Rate	0.135	0.113	0.022	0.472
Infrastructure Index	0.168	0.068	0.099***	0.001
Population Density (pp/km^2)	45.884	181.070	-135.186***	0.002
Annual Rainfall in mm	1154.432	708.184	446.248***	0.000
40-Year Average of Rainfall in mm	1163.462	712.080	451.382***	0.000
%-Deviation in Rainfall[1]	0.041	0.078	-0.037***	0.009
B. Individual Controls:				
Electricity	0.011	0.027	-0.016	0.398
Sex of the Household Head:				
Male	0.834	0.779	0.055*	0.083
Female	0.166	0.221	-0.055*	0.083
Religion:				
Orthodox	0.371	0.403	-0.032	0.816
Catholic	0.010	0.005	0.006	0.408
Protestant	0.183	0.060	0.123	0.110
Moslem	0.370	0.530	-0.160	0.269
Traditional	0.065	0.002	0.063**	0.018
Other	0.001	0.000	0.000	0.820
Age Cohorts:				
15–19	0.219	0.210	0.009	0.552
20–24	0.172	0.181	-0.009	0.550
25–29	0.175	0.165	0.011	0.528
30-34	0.140	0.130	0.010	0.317
35–39	0.136	0.131	0.004	0.706
40–44	0.076	0.095	-0.019*	0.084
45–49	0.082	0.089	-0.008	0.419
Individual Obs.	1859	2563		
Woredas	32	46		

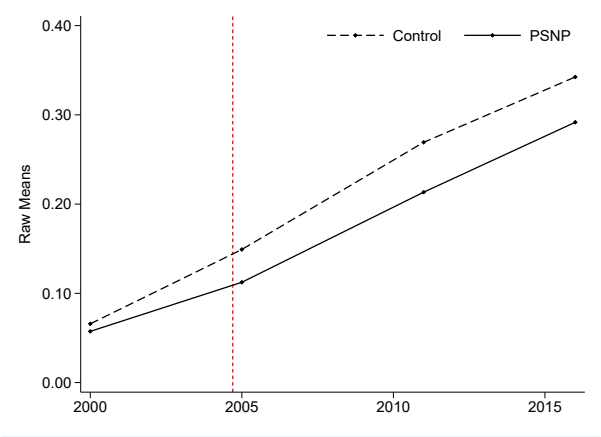
Notes: This table provides the baseline means of the control variables that are used in this paper separately for treatment and control group, as well as clustered t-tests. Means are computed from the pre-treatment sample, which is based on the DHS survey round of 2000. The sample is restricted to the rural population and to women aged 15-49. Standard errors are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

[1]The deviation in rainfall is expressed as the percentage-deviation of a 5-year average in annual rainfall from a 40-year average.

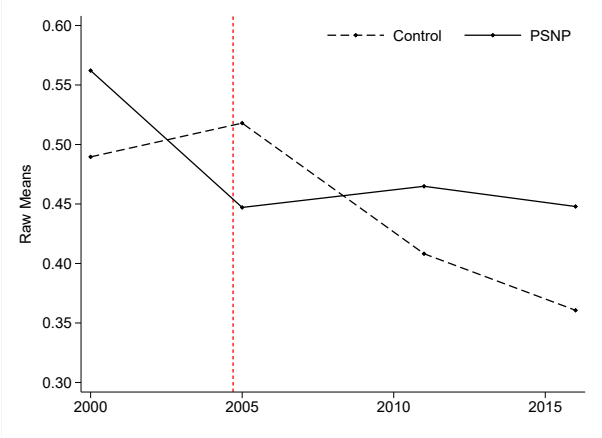
Figure 3: Raw Means of the outcomes over time



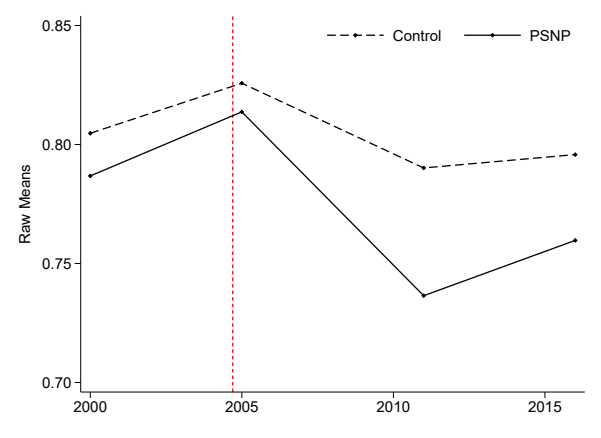
(a) Fertility



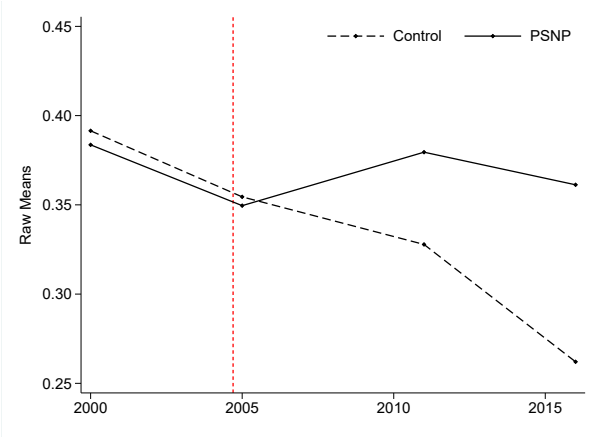
(b) Contraception



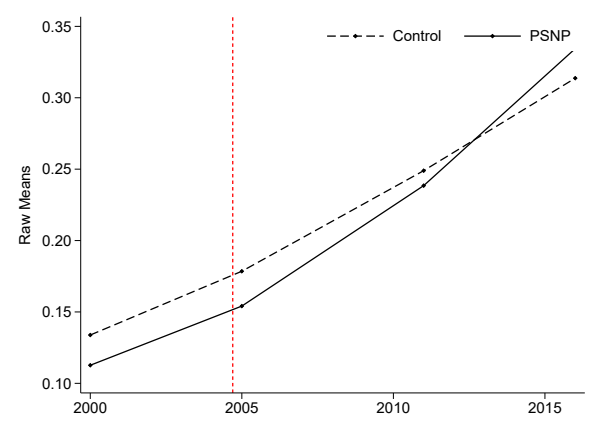
(c) Violence



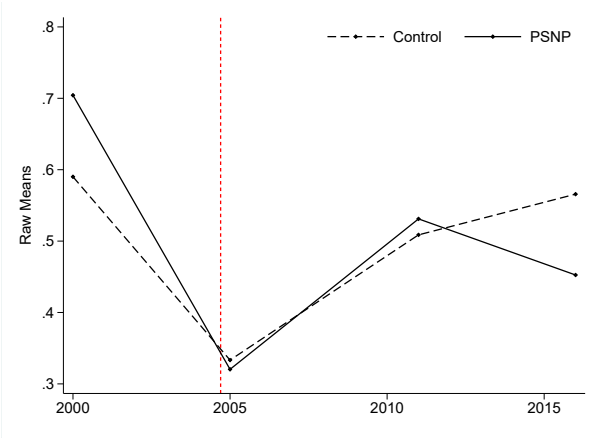
(d) Marriage



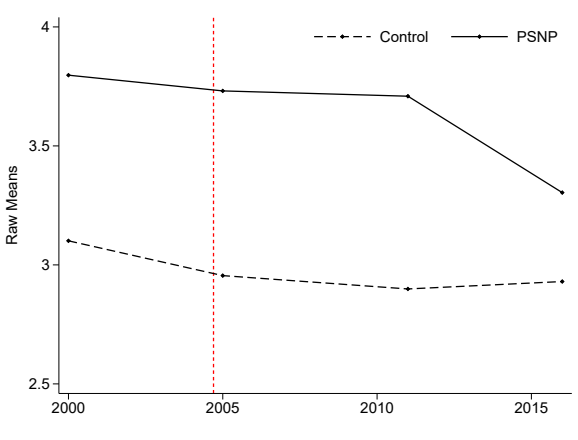
(e) Malnutrition



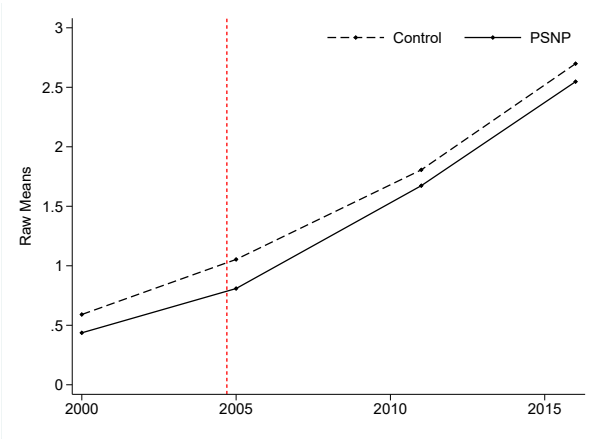
(f) Literacy



(g) Work Status



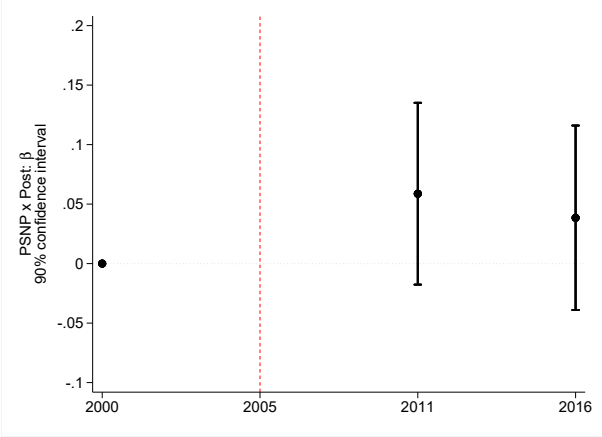
(h) Water Time



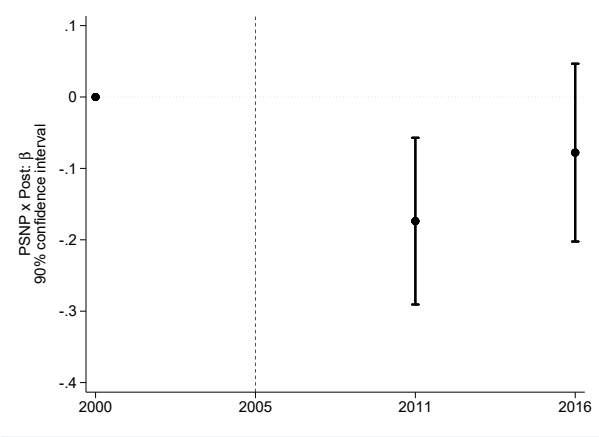
(i) Educational Attainment

Notes: This figure presents the raw means for each of the outcomes used in the analysis over time (early childbearing and general fertility are both measured by the same variable). The dashed black line shows trends of the control group, the solid black line shows the development of the treatment group. The dashed vertical red line marks the start of the PSNP programme in 2005. In this graph, the red line illustrates that the PSNP programme start preceded the DHS 2005 survey round by some months.
Source: Author's elaboration based on the Demographic and Health Surveys (DHS) 2000, 2005, 2011 and 2016.

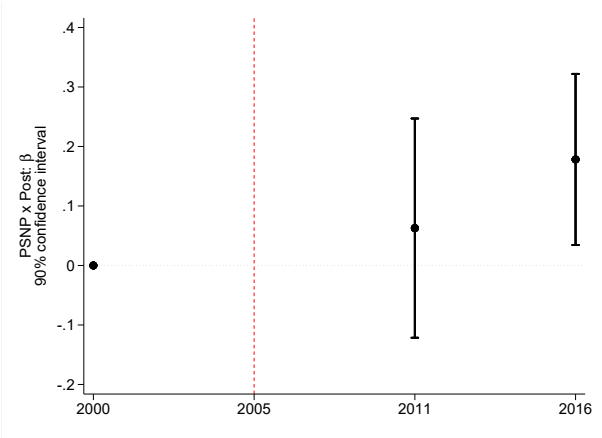
Figure 4: Event Study Graphs with single years: Main results



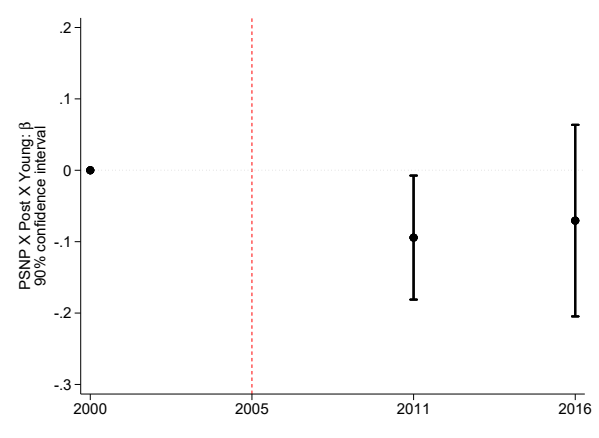
(a) Fertility



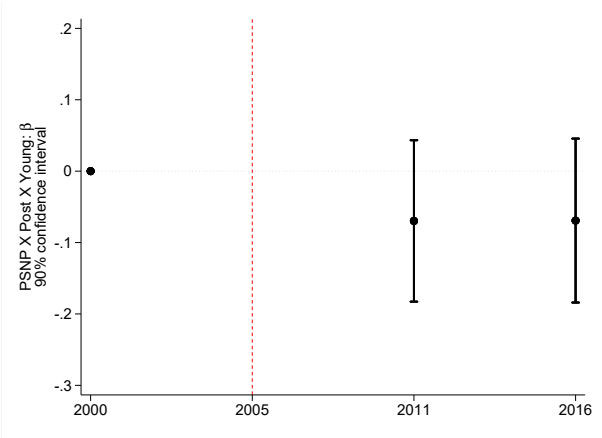
(b) Contraception



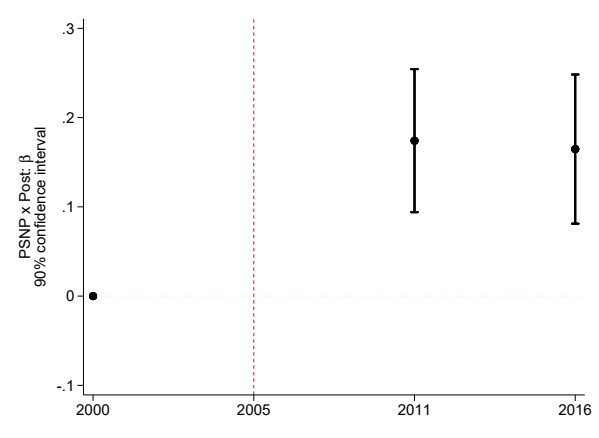
(c) Violence



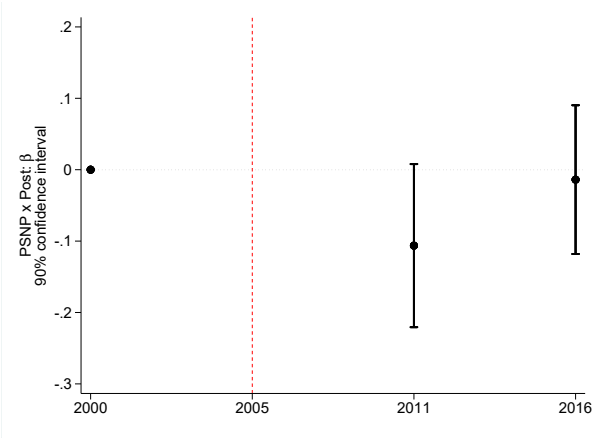
(d) Early Childbearing



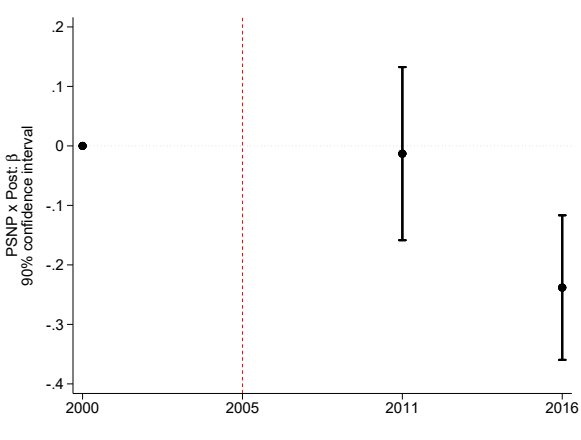
(e) Early Marriage



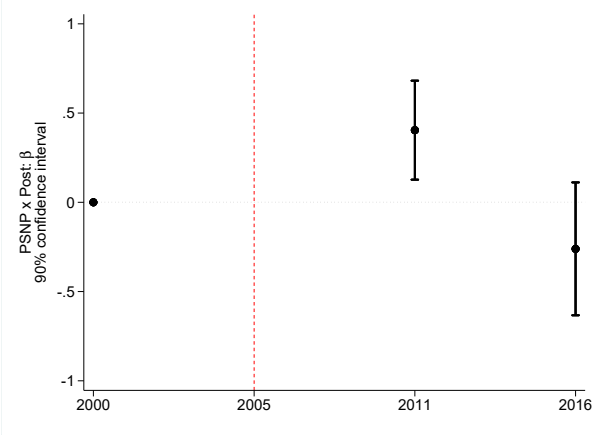
(f) Malnutrition



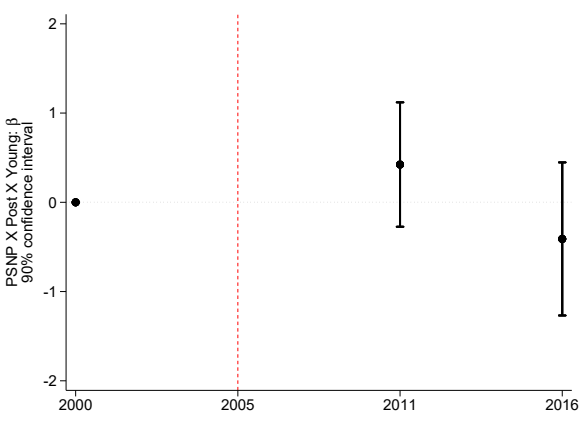
(g) Literacy



(h) Work Status



(i) Water Time



(j) Educational Attainment

Notes: This figure illustrates the regression results of Table 2 and Table 3, providing coefficients for the single years of the pooled post period. Included in this analysis are the DHS rounds of 2000, 2011 and 2016. The dashed vertical red line marks the start of the PSNP programme in 2005. The vertical black lines show the 90% confidence intervals of each coefficient. The graphs show the results for the regressions that include all control variables, which are documented in Table 2. Standard errors are clustered at Woreda-level. Reference period is the 2000 sample.
Source: Author's elaboration based on the Demographic and Health Surveys (DHS) 2000, 2011 and 2016.

Table 5: 2005 as Reference Period: Social Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Fertility	(2) Contra-ception	(3) Violence	(4) Early Childbearing	(5) Early Marriage	(6) Fertility	(7) Contra-ception	(8) Violence	(9) Early Childbearing	(10) Early Marriage
PSNP × 2000	-0.0021	0.0297	0.131*			-0.0295	0.122*	0.0181		
	(0.0342)	(0.0380)	(0.0677)			(0.0458)	(0.0693)	(0.0944)		
PSNP × Post	-0.0280	-0.0159	0.155***			0.0135	-0.0257	0.128**		
	(0.0262)	(0.0514)	(0.0380)			(0.0305)	(0.0555)	(0.0526)		
PSNP × 2000 × Young				0.0299	0.0863				0.0627	0.101*
				(0.0569)	(0.0642)				(0.0521)	(0.0530)
PSNP × Post × Young				-0.0163	9.67e-07				-0.0182	0.0312
				(0.0383)	(0.0492)				(0.0508)	(0.0581)
Observations	16,206	7,348	15,640	16,206	14,649	16,153	7,318	15,590	16,153	14,604
R^2	0.018	0.139	0.053	0.058	0.342	0.224	0.177	0.070	0.228	0.472
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

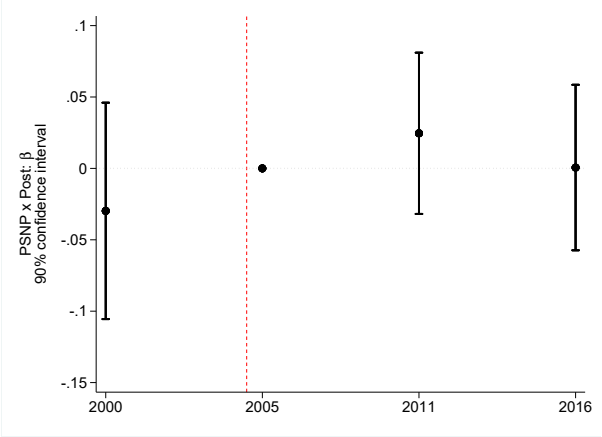
Notes: This table presents OLS regression results for variables of women's social empowerment. Here, the DHS 2005 survey is the reference period, such that row 1 and 3 show results for 2000 compared to 2005, whereas row 2 and 4 show the results for the post-period (2011 and 2016) compared to 2005. For the list of control variable and more details, see notes of Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: 2005 as Reference Period: Economic Empowerment

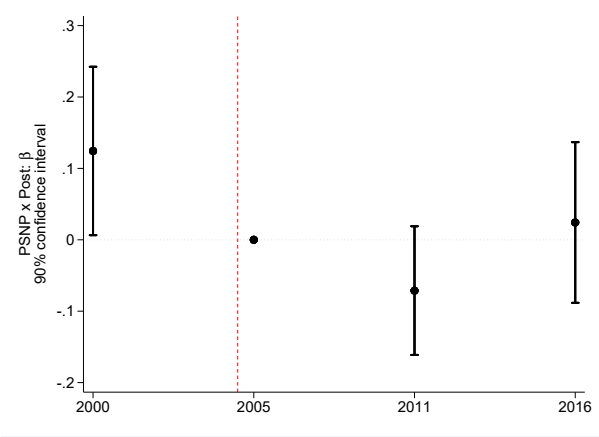
	Without Controls					Adding Woreda and Individual Controls				
	(1) Mal-nutrition	(2) Literacy	(3) Work	(4) Water Time	(5) Education	(6) Mal-nutrition	(7) Literacy	(8) Work	(9) Water Time	(10) Education
PSNP × 2000	-0.00285	-0.00325	0.136*	-0.0114		-0.114*	0.0754	0.0950	-0.135	
	(0.0433)	(0.0318)	(0.0759)	(0.160)		(0.0580)	(0.0484)	(0.0935)	(0.264)	
PSNP × Post	0.0713**	0.0242	-0.0188	-0.0972		0.0692*	0.00610	-0.0603	-0.000497	
	(0.0351)	(0.0355)	(0.0587)	(0.154)		(0.0379)	(0.0415)	(0.0704)	(0.173)	
PSNP × 2000 × Young					-0.0353					0.0369
					(0.256)					(0.347)
PSNP × Post × Young					0.256					0.0707
					(0.373)					(0.354)
Observations	12,185	15,898	15,985	16,155	16,206	12,136	15,850	15,932	16,106	16,153
R^2	0.040	0.099	0.195	0.215	0.249	0.061	0.237	0.221	0.263	0.340
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents OLS regression results for variables of women's economic empowerment. For details see notes of Table 5. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

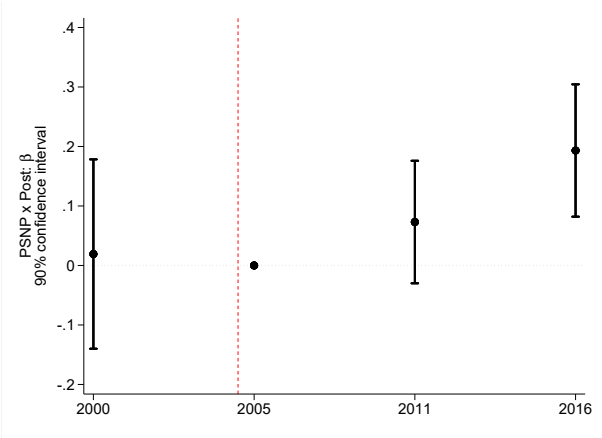
Figure 5: Event Study Graphs with single years: 2005 as Reference Period



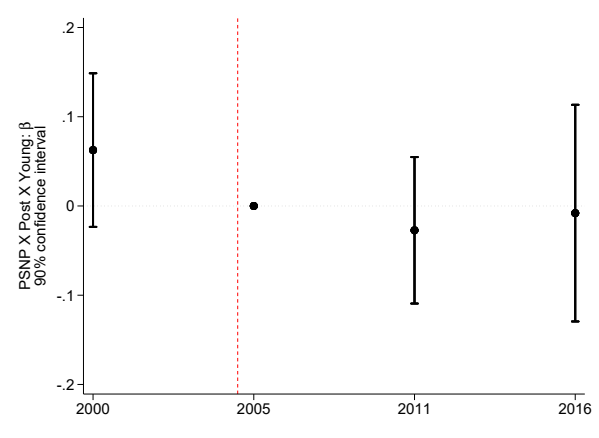
(a) Fertility



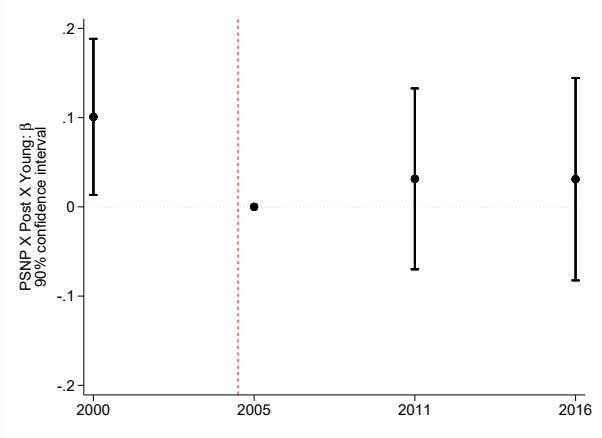
(b) Contraception



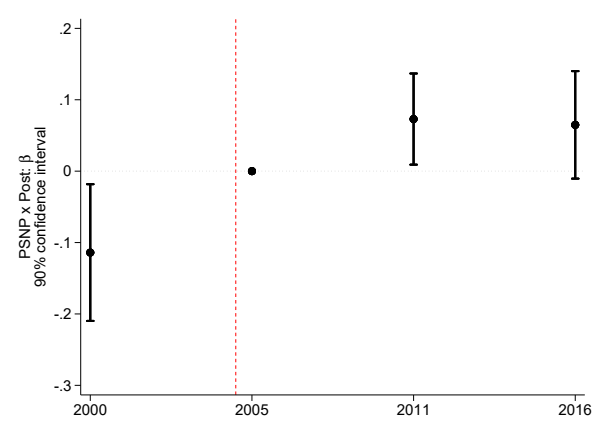
(c) Violence



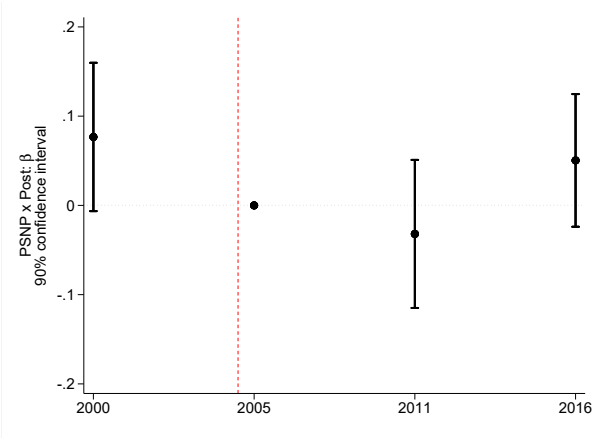
(d) Early Childbearing



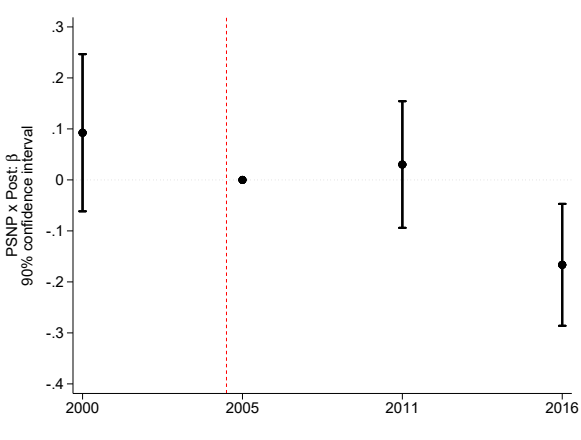
(e) Early Marriage



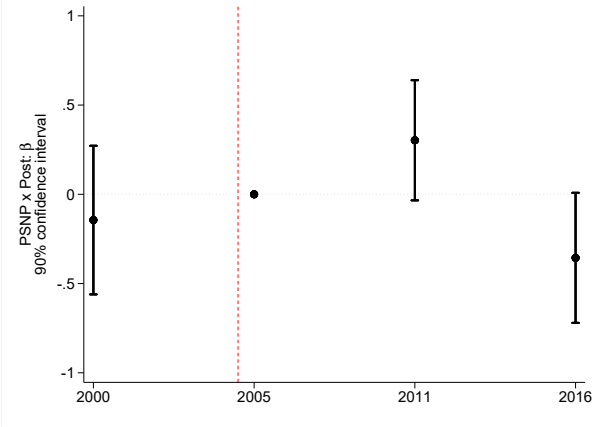
(f) Malnutrition



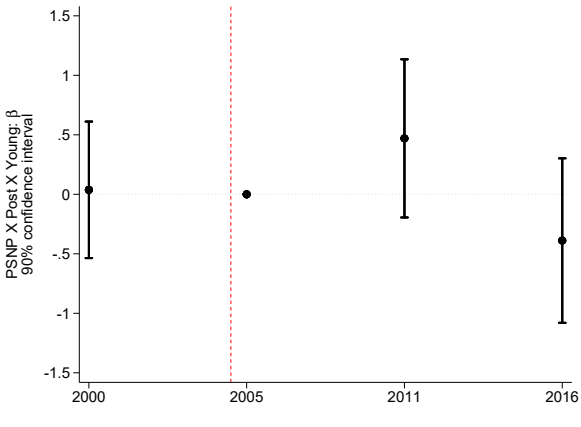
(g) Literacy



(h) Work Status



(i) Water Time



(j) Educational Attainment

Notes: This figure illustrates the regression results of Table 5 and Table 6, providing coefficients for the single years of the pooled post period. Included in this analysis are the DHS rounds of 2000, 2005, 2011 and 2016. The dashed vertical red line marks the start of the PSNP programme in 2005. The slightly shifted position of the red line illustrates that the programme started some months before the 2005 DHS survey round. The vertical black lines show the 90% confidence intervals of each coefficient. The graphs show the results for the regressions that include all control variables, which are documented in Table 2. Standard errors are clustered at Woreda-level. Reference period is the 2005 sample.
Source: Author's elaboration based on the Demographic and Health Surveys (DHS) 2000, 2005, 2011 and 2016.

Table 7: Continuous Treatment – Social Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Fertility	(2) Contra-ception	(3) Violence	(4) Early Childbearing	(5) Early Marriage	(6) Fertility	(7) Contra-ception	(8) Violence	(9) Early Childbearing	(10) Early Marriage
PSNP (Intensity) × Post	-0.0036	-0.0262**	-0.0154			0.0141	-0.053***	-0.0189		
	(0.0109)	(0.0119)	(0.0218)			(0.0108)	(0.0143)	(0.0261)		
PSNP (Intensity) × Post × Young				-0.0160	-0.0283				-0.0177	-0.0216
				(0.0162)	(0.0196)				(0.0159)	(0.0175)
Observations	12,633	6,414	12,246	12,633	11,411	12,584	6,387	12,200	12,584	11,370
R^2	0.020	0.142	0.063	0.065	0.354	0.229	0.183	0.075	0.232	0.481
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's social empowerment, using a continuous treatment variable. The treatment variable is a continuous measure of the ordinal treatment intensity from 0 to 4. Results present the effect of an increase in treatment intensity by one level. The analysis is based on the DHS survey rounds of 2000, 2011 and 2016. For the list of control variables and more details, see notes of Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Most Treated Woredas Sample – Social Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Fertility	(2) Contra-ception	(3) Violence	(4) Early Childbearing	(5) Early Marriage	(6) Fertility	(7) Contra-ception	(8) Violence	(9) Early Childbearing	(10) Early Marriage
PSNP × Post	-0.0087	-0.100**	-0.0437			0.0539	-0.223***	0.0790		
	(0.0385)	(0.0437)	(0.0717)			(0.0434)	(0.0516)	(0.0783)		
PSNP × Post × Young				-0.0778	-0.115*				-0.0759	-0.0540
				(0.0593)	(0.0673)				(0.0605)	(0.0630)
Observations	9,912	4,964	9,617	9,912	8,955	9,869	4,939	9,577	9,869	8,919
R^2	0.019	0.144	0.074	0.064	0.368	0.224	0.187	0.094	0.228	0.493
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's social empowerment, restricting the sample to the most treated Woredas. Here, the treatment group comprises only Woredas that are treated with an intensity level of 26%–42% or 43%–65%. The analysis is based on the DHS survey rounds of 2000, 2011 and 2016. For the list of control variables and more details, see notes of Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Continuous Treatment – Economic Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Mal-nutrition	(2) Literacy	(3) Work	(4) Water Time	(5) Education	(6) Mal-nutrition	(7) Literacy	(8) Work	(9) Water Time	(10) Education
PSNP (Intensity) × Post	0.0261***	0.0111	-0.0491*	0.0209		0.0342***	-0.0181	-0.0422	0.103*	
	(0.00967)	(0.0129)	(0.0258)	(0.0549)		(0.0120)	(0.0156)	(0.0262)	(0.0600)	
PSNP (Intensity) × Post × Young					0.182					0.119
					(0.110)					(0.106)
Observations	10,738	12,371	12,630	12,613	12,633	10,692	12,326	12,581	12,568	12,584
R^2	0.046	0.105	0.189	0.203	0.246	0.065	0.247	0.215	0.264	0.348
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's economic empowerment, using a continuous treatment variable. See notes of Table 7 for more details. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Most Treated Woredas Sample – Economic Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Mal-nutrition	(2) Literacy	(3) Work	(4) Water Time	(5) Education	(6) Mal-nutrition	(7) Literacy	(8) Work	(9) Water Time	(10) Education
PSNP × Post	0.101***	0.0372	-0.179**	0.105		0.178***	-0.0883	-0.134	0.107	
	(0.0348)	(0.0441)	(0.0875)	(0.166)		(0.0474)	(0.0550)	(0.0844)	(0.142)	
PSNP × Post × Young					0.602					0.0830
					(0.397)					(0.440)
Observations	8,465	9,683	9,909	9,897	9,912	8,425	9,643	9,866	9,855	9,869
R^2	0.046	0.112	0.201	0.234	0.271	0.064	0.265	0.230	0.277	0.372
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the main OLS regression results for variables of women's economic empowerment, restricting the sample to the most treated Woredas. See notes of Table 8 for more details. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Probit Model: Social Empowerment

	Without Controls					Adding Woreda and Individual Controls				
	(1) Fertility	(2) Contra-ception	(3) Violence	(4) Early Childbearing	(5) Early Marriage	(6) Fertility	(7) Contra-ception	(8) Violence	(9) Early Childbearing	(10) Early Marriage
PSNP × Post	-0.0572	-0.135	0.0775			0.117	-0.748**	0.335		
	(0.0912)	(0.181)	(0.203)			(0.134)	(0.304)	(0.237)		
PSNP × Post × Young				-0.144	-0.430**				-0.285	-0.272
				(0.162)	(0.185)				(0.196)	(0.238)
Observations	12,633	6,260	12,246	12,633	11,411	12,584	6,233	12,200	12,584	11,368
Pseudo R^2	0.0145	0.151	0.0462	0.0496	0.341	0.187	0.195	0.0567	0.190	0.465
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	×	✓	✓	✓	✓	✓
Individual Controls	×	×	×	×	×	✓	✓	✓	✓	✓

Notes: This table presents the results of the Probit model for the binary variables of women's social empowerment, resulting from Maximum Likelihood Estimations. The analysis is based on the DHS survey rounds of 2000, 2011 and 2016. For the list of controls and more details, see Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Probit Model: Economic Empowerment

	Without Controls			Adding Woreda and Individual Controls		
	(1) Malnutrition	(2) Literacy	(3) Work	(4) Malnutrition	(5) Literacy	(6) Work
PSNP × Post	0.211**	0.151	-0.443**	0.475***	-0.552***	-0.345
	(0.0930)	(0.118)	(0.188)	(0.130)	(0.201)	(0.243)
Observations	10,738	12,371	12,630	10,692	12,324	12,581
Pseudo R^2	0.0348	0.110	0.151	0.0516	0.269	0.173
Woreda FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	✓	✓	✓
Individual Controls	×	×	×	✓	✓	✓

Notes: This table presents the results of the Probit model for all binary variables of women's economic empowerment, similarly to Table 11. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Heterogeneity in Employment Effects

	Without Controls				Adding Woreda and Individual Controls			
	(1) Work$_{\geq 18}$	(2) Work (only paid)	(3) Paid Work	(4) Work for Others	(5) Work$_{\geq 18}$	(6) Work (only paid)	(7) Paid Work	(8) Work for Others
PSNP × Post	-0.147**	-0.133	0.0751	0.0851**	-0.0930	-0.0342	0.232	0.0704*
	(0.0666)	(0.0837)	(0.0917)	(0.0327)	(0.0765)	(0.0826)	(0.140)	(0.0386)
Observations	10,771	9,732	7,094	7,093	10,726	9,690	7,065	7,064
R^2	0.206	0.177	0.148	0.116	0.229	0.230	0.207	0.134
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	×	×	×	×	✓	✓	✓	✓
Individual Controls	×	×	×	×	✓	✓	✓	✓

Notes: This table presents OLS regressions for variations in women's work status. Column 1 shows results for the work variable that excludes underage women from the sample, the work definition in column 2 includes all women from 15–49, but excludes unpaid work from the definition. Column 3 shows results for being paid for one's work compared to not getting paid, column 4 the results for the likelihood to have worked for others, compared to being self-employed or working for family members. Column 3 and 4 are restricted to currently working women. Column 5 to 8 show results of the same analysis once controls are included. The analysis is based on the DHS rounds 2000, 2011 and 2016. For more details see notes of Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Heterogeneity in Employment Effects: Young Women

	Young Cohort 18–22					Young Cohort 18–24				
	(1) Work	(2) Work$_{\geq 18}$	(3) Work (only paid)	(4) Paid Work	(5) Work for Others	(6) Work	(7) Work$_{\geq 18}$	(8) Work (only paid)	(9) Paid Work	(10) Work for Others
PSNP × Post × Young$_{18-22}$	0.107*	0.113*	0.132**	-0.0123	-0.0730					
	(0.0552)	(0.0592)	(0.0566)	(0.101)	(0.0502)					
PSNP × Post × Young$_{18-24}$						0.112**	0.126**	0.152***	-0.0314	-0.0150
						(0.0523)	(0.0549)	(0.0483)	(0.0904)	(0.0421)
Observations	12,581	10,726	9,690	7,065	7,064	12,581	10,726	9,690	7,065	7,064
R^2	0.214	0.230	0.231	0.210	0.137	0.215	0.230	0.233	0.212	0.136
Woreda FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Woreda Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table presents OLS regression results for the same variables as Table 13, now using a triple differences framework to test work effects for young women. In column 1 the work definition that was used throughout the main analysis is added compared to Table 13. Row 1 shows the results for women aged 18 to 22, row 2 presents results for women aged 18–24. The analysis is based on the DHS rounds 2000, 2011 and 2016. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Heterogeneity in Early Childbearing and Marriage

	Cohort 15–20		Cohort 15–24		Cohort 18-22	
	(1) Early Childbearing	(2) Early Marriage	(3) Early Childbearing	(4) Early Marriage	(5) Early Childbearing	(6) Early Marriage
PSNP × Post × $Young_{15-20}$	-0.0816	-0.0554				
	(0.0587)	(0.0644)				
PSNP × Post × $Young_{15-24}$			-0.0613	-0.0448		
			(0.0612)	(0.0586)		
PSNP × Post × $Young_{18-22}$					-0.0912	-0.0431
					(0.0690)	(0.0766)
Observations	12,584	11,370	12,584	11,370	12,584	11,370
R^2	0.232	0.481	0.232	0.479	0.234	0.493
Woreda FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Woreda Controls	✓	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓	✓

Notes: This table presents OLS regression results for effects on early childbearing in a triple differences framework, with varying definitions of the young cohort. Column 1 and 2 compares effects for 15-20 year old women to the rest, columns 3-4 and 5-6 do the same for women aged 15-24 and 18-22, respectively. The analysis is based on the DHS rounds 2000, 2011 and 2016. For more details and the list of controls see Table 2. Standard errors in parenthesis are clustered at Woreda-level. *** p<0.01, ** p<0.05, * p<0.1.

Data Sources and Construction

Panel Construction: The primary data source for my analysis are four rounds of the Demographic and Health surveys (DHS) of the years 2000, 2005, 2011 and 2016, whereby I rely on the women's recode data sets. My identification strategy consists in comparing changes in Woredas over time, while the DHS data identifies single observations only at the regional level as lowest administrative unit. Therefore, I used the DHS geographic data sets that are provided for each survey round and that contain information on latitude, longitude and elevation of each survey cluster for the use in Geographic Information Systems (GIS) (Croft et al. 2018). Using GIS software (QGIS) and an administrative shapefile of Ethiopia at Woreda-level from 2013, provided by Open Africa (2021), I matched DHS survey clusters with the respective Woredas. Merging the joined file with the initial DHS recode data sets, I assigned a Woreda to each DHS cluster and was thus able to identify individual observations at Woreda-level[5]. To construct the panel that I used for the main analyses in this paper, I asssigned each Woreda a treatment status (as explained below) and appended the different DHS survey rounds. Moreover, I dropped all urban observations and non-permanent residents of a household, to reduce bias. I also balanced the panel to include only the Woredas that were sampled across all 4 DHS survey rounds. This caused a considerable reduction in the size of control and treatment group, since the DHS did not cover all/the same Woredas in each round. Finally, I added the district-level controls like population density and rainfall to the panel.

Treatment Assignment: Precise official information on which Woredas have been covered by the PSNP programme over time was challenging to obtain. To identify targeted Woredas, I therefore relied on two maps from UNOCHA (2007) and World Bank (2010). These maps are accurate enough to identify treated Woredas and also allow to observe possible changes in roll-out between 2007 and 2009.
One issue with the identification of the treatment group was the administrative reform in Ethiopia that has been mentioned in chapter 2. The reform caused changes in boundaries and names of Woredas, mostly between 2006 and 2008 and it almost exclusively increased the number of Woredas through territorial divisions. Therefore, I compared the two maps of 2007 and 2009 with an administrative shapefile of 2013[6] and assigned a Woreda with the name as indicated in the 2013-administrative shapefile to the treatment group, when it was marked as treated in either of the maps (however, I rely on the 2009 map most of the time, since it is more up to date and reflects the areas after the reform). The shapefile almost exactly matches the administrative boundaries of the 2009 map and I also used it to assign Woreda names to DHS clusters. Using the same shapefile to define the treatment group and assign Woreda names to DHS observations, I can thus guarantee that every DHS cluster is assigned the correct treatment or control Woreda and changes in names or boundaries are not an issue. There were only two small remaining issues: Firstly, some of the Woredas were further divided between 2009 and 2013. In this case, I decided to assign

[5]The administrative reforms that caused changes of Woreda boundaries and names over time did not pose a problem here, since the geocoding allowed me to assign the same Woreda name with the same boundaries to cluster geocoordinates of different rounds.

[6]I do not use a newer map, since boundaries and names were subject to further changes after 2013. Also, I did not find a shapefile of 2009, so the 2013 map is closest to the Worldbank map of treatment assignment

all the Woredas of the 2013 map, that were on an area marked as treated in the 2009 map, to the treatment group. Secondly, there were a few cases of Woredas that appeared as targeted in the 2007 UN-map, but not in the 2009 map. I decided to include them in the treatment group, since it is likely that they were targeted by the PSNP at least for some years. However, in both cases this concerns only a small number of Woredas and they do not appear in the final panel.

Outcomes

The main outcomes of interest are exclusively taken from the Demographic and Health Surveys. I restrict my analysis to the individual recode data set of the DHS for each round, which includes only interviewed women of age 15 to 49. Detailed information on all variables of the data set are provided in the respective DHS recode manual (CSA 2013) and detailed information on the questionnaires of the survey is taken from the final reports (CSA 2001). Based on the data sets, I construct the following outcomes for the analysis:

Fertility/Early Childbearing: These variables are based on the number of births that sampled women have given in the three years preceding the interview. In the data, the categorical variable "total number of births in the last three years" ranges from 0 to 3 births. For measurement purposes, and since observations with a number of births greater than one are very rare, I create the binary dummy variable "has given birth in the last 3 years" that equals one, if women have given birth to 1 or more children in the 3 years preceding the survey, and zero if they have not given birth.

Marriage: This variable is constructed using the categorical variable "current marital status of the respondent" in the surveys. The variable in the DHS samples comprises six different categories: "Married", "never married", "Living together", "widowed", "divorced" and "not living together (anymore)". "Never married" can include women who have a partner, but are not married and do not live together with him/her. I construct the binary marriage variable "married or cohabiting" that equals one, if women are "married" or "living together" and zero if they are "not married" or "not living together". I exclude divorced and widowed women from the analysis.

Attitude towards sexual violence (Violence): This binary variable is based on the yes-no question "In your opinion, is a husband justified in hitting or beating his wife if she refuses to have sex with him?" in the DHS survey. The binary variable shall reflect women's attitude towards sexual violence, her control over the body and general social norms. Here, all women are excluded from the analysis who answered "don't know".

Contraception: The variable is based on the categorical variable "Current contraceptive method" in the DHS samples. It lists more than 10 different traditional, folkloric and modern contraceptive methods, including pill, IUD, injections, condom or female sterilization. I define using contraceptives as relying on

any contraceptive method. Traditional methods like withdrawal and periodic abstinence might be less reliable, however, they still show the willingness to prevent childbearing and they constitute only a very small number of observations in the panel. The binary "contraception" variable thus equals one, if a respondent is currently using any of the methods, and zero if she is not using any method. To reduce bias, I restrict the sample for this variable to women who were sexually active in the four weeks preceding the interview.

Malnutrition: I measure malnutrition using the Body Mass Index (BMI) of interviewed women. Thereby, I use the UN cutoff point for mild malnutrition that has been recommended for indicating chronic energy deficiency among non-pregnant women, such that I consider all women to be (mildly, moderately or severely) malnourished whose BMI is below 18.5 (WFP 2005). Following UN and DHS recommendations of measuring the BMI, I exclude women who were pregnant at the time of the interview or had given birth in the two months preceding the interview from the sample, to avoid bias.

Education: Educational attainment of young women is measured by a continuous variable that indicates the overall education in single years that respondents attained. It is constructed from the highest educational level achieved (including: no education, primary, secondary, higher) and the grade at that level, and ranges from zero to 17 years in my panel.

Literacy: Literacy is based on a question to respondents who attended primary schooling whether they can read a whole or part of a sentence shown in the interview. I consider all respondents to be illiterate who were not able to read the sentence at all, and literate when they were able to either completely, or partly read the sentence. Respondents who attended at least secondary education are classified as literate.

Work: To measure the work status of women, I use the categorical variable "Whether the respondent worked in the last 12 months" of the DHS surveys. It includes the categories "no", "currently working", "in the past year" and "on leave". I aggregate the categories to the binary variable "work", which equals one if a woman either worked in the past 12 months or is currently working (or if she is on leave), and zero otherwise. Since Public Works opportunities of the PSNP are seasonal, it is of interest to include the share of women who worked in the past year.
For the robustness checks I use two additional variables. **Paid Work** is based on a question whether the respondent is paid for her work. I code it to equal one, if the respondent is either paid in cash, in kind, or both, and zero if the respondent is not paid for her work. Furthermore, **Work for others** is based on the question whether a respondent works for a family member, is self-employed or works for others. I construct a binary variable that equals one if the respondent worked for others and zero if she is self-employed or works for family members. Both variables are based on respondents who are currently working or worked in the past 12 months.

Time to Get Water: The variable excludes all respondents whose household has piped drinking water or use rainwater or bottled water and is based on the question "How long does it take you to go there (i.e. the main source of drinking water), get water and come back?" (in minutes). I drop all not de jure residents, and recode all answers that indicated that the water source is directly available in the household ("On premises") as one minute, and all answers that indicated to need more than one day to collect water as 1440 minutes (minutes of one whole day). In my analysis, the "Water Time" variable gives the logarithmized time in minutes that respondents need to get to the next drinking water source and back.

Control Variables

Individual Controls:

Individual controls are taken from the respective DHS women's recode data sets and comprise four different variables. **Electricity** is a dummy variable for whether the respondent's household has electricity or not. Reference category is that the household does not have electricity. **Religion** is a categorical variable, including "orthodox", "catholic", "protestant", "moslem", "traditional" and "other". Reference category is "orthodox". The **age-group** variable groups respondents into 7 different categories of age: 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49. Reference category are 15-19 year old women. Finally, **sex of the household head** is a binary variable with the reference category being male.

Woreda-level controls:

Population Density: Data for the population densities are taken from the Gridded Population of the World data set of the NASA Socioeconomic Data and Applications Center (SEDAC) and the CIESIN at Columbia University (CIESIN 2017). The "Adminitrative Unit Center Points" data set provides UN WPP-adjusted population estimates by administrative unit center point location. For Ethiopia, estimates are based on the census of 2007 (CIESIN 2018). I imported the population estimates as shapefile and matched the density unit points with the administrative shapefile for Ethiopia (of 2013) that is used throughout the thesis, using GIS software. Thus, I could assign population densities to each Woreda and merge them with my panel. Densities are not provided separately for rural and urban areas in most cases. In the few cases where both rural and urban densities were provided for one Woreda, I only relied on the rural density figure.

Rainfall: I draw on rainfall data from the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) database of the Climate Hazard Center (UC Santa Barbara) and the US Geological Survey (USGS) (CHC 2021), using the Early Warning Explorer (EWX). The EWX was developed to address data limitations in climate-risk vulnerable regions and to monitor extreme events like droughts, serving as early warning tool for food insecurity (Anthony et al. 2021). To obtain Woreda-level figures of annual rainfall, I pulled the annual pentadal precipitation (mm) time-series from 1981–2020 with the EWX for each Woreda and merged them into one data set. In the few cases where the EWX Woredas had changed names or borders compared to the map of 2013, I used the data from the Woreda whose borders were

closest to that of the shapefile of 2013.

The long-term average of rainfall is computed as mean from 1981 to 2020. I construct my rainfall deviation variable by computing the annual average in rainfall for the 5 years preceding each of the 4 survey rounds (including the survey year) and taking its deviation from the 40-year mean.

Poverty Rate, Literacy Rate, Infrastructure Index:

These pre-treatment district-level controls are based on the DHS survey of 2000, whereby single observations are aggregated to the Woreda-level for each district. As in the main analysis, figures for these variables are restricted to the rural population. The **Literacy Rate** is constructed by aggregating the variable for literacy (see above) at Woreda-level.

I construct a **Poverty Rate** as measure for the relative wealth distribution within a Woreda. The variable is based on the wealth index of the DHS data sets, which is a composite measure of a household's living standard. It is constructed by a range of various indicators like sanitation facilities, housing materials or a household's assets and ownerships.The index places households on a continuous scale of relative wealth, separating them into five wealth quintiles (CSA 2013). From those quintiles I construct a "poverty" variable that equals one, if the respondent's household belongs to the two lowest wealth quintiles, and zero if the household belongs to the three highest wealth quintiles.

The **Infrastructure Index** is based on two DHS survey questions: First, whether the respondent's household has electricity or not, and second, how much time the respondent needs to get to the next water source to collect water, and back. From the question for the water time, I construct a dummy that equals one if a respondent needs less than 15 minutes to the next water source and back. The Infrastructure Index is then constructed as simple cumulative index of these two binary variables, normalized to one. Thus, it equals one when both the respondent's household has electricity and the respondent needs less than 15 minutes to collect water.