ETHIOPIA
Strategy Support Program

ETHIOPIAN DEVELOPMENT RESEARCH INSTITUTE
EDRI

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
IFPRI
A member of the CGIAR Consortium

ESSP WORKING PAPER 99

The impact of Ethiopia's Productive Safety Net Programme on the nutritional status of children: 2008–2012

Guush Berhane, John Hoddinott, and Neha Kumar

January 2017

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Ethiopia's Productive Safety Net Programme (PSNP) is a large-scale social protection intervention aimed at improving food security and stabilizing asset levels. The PSNP contains a mix of public works employment and unconditional cash and food transfers. It is a well-targeted program; however, several years passed before payment levels reached the intended amounts. The PSNP has been successful in improving household food security. However, children's nutritional status in the localities where the PSNP operates is poor, with 48 percent of children stunted in 2012. This leads to the question of whether the PSNP could improve child nutrition. In this paper, we examine the impact of the PSNP on children's nutritional status over the period 2008–2012. Doing so requires paying particular attention to the targeting of the PSNP and how payment levels have evolved over time. Using inverse-probability-weighted regression-adjustment estimators, we find no evidence that the PSNP reduces either chronic undernutrition (height-for-age z-scores, stunting) or acute undernutrition (weight-for-height z-scores, wasting). While we cannot definitively identify the reason for this non-result, we note that child diet quality is poor. We find no evidence that the PSNP improves child consumption of pulses, oils, fruits, vegetables, dairy products, or animal-source proteins. Most mothers have not had contact with health extension workers nor have they received information on good feeding practices. Water practices, as captured by the likelihood that mothers boil drinking water, are poor. These findings, along with work by other researchers, have informed revisions to the PSNP. Future research will assess whether these revisions have led to improvements in the diets and anthropometric status of preschool children in Ethiopia.

Keywords: nutrition, stunting, social protection, Ethiopia

1. INTRODUCTION

There is increasing interest in policies and interventions that improve the nutritional status of preschool children. This interest reflects two considerations. First, improvements in nutritional status are an intrinsically valuable development outcome. Second, the preponderance of evidence shows that the harm caused by undernutrition in early life—both lost physical growth and neurological damage—is not fully recovered, leading to lower levels of height, schooling, cognitive skills, and ultimately income in adulthood (Black et al. 2013; Hoddinott et al. 2013).

Black et al. (2013) identify two sets of interventions that have potential to improve nutrition: nutrition-specific interventions, which address the immediate causes of poor nutritional status, and nutrition-sensitive interventions, which address the underlying determinants of malnutrition. Within this context, examining the impact of Ethiopia's Productive Safety Net Programme (PSNP) on the nutritional status of children, is of particular interest. The program has some of the characteristics of a nutrition-sensitive intervention—it is targeted toward a food-insecure population and, through its provision of cash and food transfers, addresses one of the underlying determinants of undernutrition. It is one of the largest safety net programs in Africa south of the Sahara, thus allowing us to see whether such interventions can have impacts when implemented at scale. Moreover, Ethiopia is an instructive country to study because undernutrition, while declining, remains prevalent and widespread, and there is considerable government commitment to address the problem (Lemma and Matji 2013).

Our assessment of the impact of the PSNP on children's nutritional status over the period 2008–2012 begins with a description of the quantitative data available to us. We then outline the PSNP in detail, explaining the objectives of the program and how it has been implemented. We pay particular attention to the targeting of the PSNP and how payment levels have evolved over time. We then discuss issues associated with estimating impact in the context of the implementation of the PSNP before describing the impact estimator we use, inverse-probability-weighted regression adjustment (IPWRA). After explaining how we implement this estimator, we present our estimates of the impact of the PSNP on the nutritional status (height and weight) of preschool children. We describe these data and show trends over time in chronic (height-for-age z-scores, stunting) and acute (weight-for-height z-scores, wasting) malnutrition. Strikingly, we find no evidence of impact over the period 2008–2012. In light of these findings, we explore some of the possible reasons for this outcome, focusing on mothers' lack of contact with health and nutrition services, children's low levels of consumption of non-staple foods, and poor water and sanitation practices.

2. DATA

We draw on four rounds of data. Starting in 2006, we collaborated with the government of Ethiopia's Food Security Coordination Directorate (which implements the PSNP) and Ethiopia's Central Statistical Agency (CSA) on the design and implementation of a longitudinal survey of PSNP beneficiaries and non-beneficiaries. Based on the list of *woredas*[1] initially included in PSNP, 68 woredas were randomly sampled on a probability proportional to size basis, stratified by region. Within each woreda, sample enumeration areas (EAs) were randomly selected from a list of EAs with PSNP activities. Within each EA, a list of households was constructed that included information on PSNP beneficiary status. From these lists, 15 PSNP beneficiary and 10 non-beneficiary households were randomly sampled. These sample sizes were based on power calculations showing how large the sample needed to be to identify an effect size equivalent to a 10-percentage-point increase in food security, as assessing whether the PSNP was improving household food security was the original objective of the survey. The first survey round was fielded in July and early August 2006; a second round between late May and early July 2008; a third in June and July 2010; and the fourth in June and July 2012.

Interviews with sample households collected information on household demographic composition, assets, agriculture, non-agricultural income-generating activities, consumption, food security, and shocks. This longitudinal survey contains extensive information on PSNP participation of households, including the months and years in which the household undertakes public works, the duration of participation, the number of days worked by all household members (including all men and all women), and payment received for this work. The surveys also include detailed information on targeting criteria and the correlates of program participation, including pre-program household and locality characteristics. Anthropometric information was collected in 2008, 2010, and 2012, but not 2006. Household-level data collection was complemented by a quantitative community survey that included modules on access to health facilities and a community price questionnaire.

The 2006 survey generated data on 3,680 households in 148 EAs within 68 woredas. In 2008, the CSA enumerator assigned to each EA was provided with the list of households interviewed during 2006, and used these data along with, in some cases, maps and assistance from local officials and residents in locating households for re-interview. Attrition was low. Only 137 households (or 3.7 percent) of the baseline sample were not re-interviewed during the 2008 survey. About a third of these households are from two EAs within Oromiya, where the survey could not take place. The 2008 resurvey covered 3,527 households in 146 EAs within 68 woredas. In 2008, CSA also surveyed households in the Amhara region, where the PSNP has been supported by the United States Agency for International Development (USAID) through grants to eight nongovernmental organizations. These households are referred to as the Amhara-HVFB sample.[2] Data were collected from 1,163 households within 44 *kebeles*[3] located in 11 woredas. In total, the 2008 survey generated data on 4,4690 households. In 2010, households in the 2006–2008 panel were traced and re-interviewed, as were the households in Amhara-HVFB; 4,645 households were interviewed.

The government of Ethiopia requested that the 2012 survey be expanded to include households that had been recently graduated from the PSNP as well as households that received direct support payments (unconditional transfers given largely to elderly households or households containing disabled individuals; see Section 3 for further details) using lists of these groups held at the kebele level, so as to ensure adequate representation of these types of households. Consequently, in 2012, 5,092 households were interviewed. Table 2.1 gives the number of households interviewed by round and region.

Table 2.1 Number of households interviewed, by round and region

	Tigray	Amhara	Amhara-HVFB	Oromiya	SNNPR	Total
2006	897	894	-	939	950	3,680
2008	868	867	1,163	861	931	4,690
2010	846	847	1,150	885	917	4,645
2012	991	985	1,103	965	1,048	5,092

Source: Authors' calculations.
Note: HVFB = high-value food basket; SNNPR = Southern Nations, Nationalities, and Peoples' Region.

[1] A *woreda* is an administrative unit equivalent to a district or county.

[2] HVFB refers to the "high-value food basket" that these households receive.

[3] A *kebele* is an administrative sub-unit of a woreda. It is equivalent to a sub-district.

Attrition in the sample is low, especially considering the physical inaccessibility of many of the survey localities and the fact that this was the first longitudinal survey ever conducted by CSA. In the non-Amhara-HVFB localities, out of the 3,680 households interviewed in 2006, 3,197 were interviewed in 2012. Out of the 483 households that attrited, 101 were lost because the kebele was not resurveyed. Between 2006 and 2012, attrition was 13.1 percent, or 2.1 percent per year. For the Amhara-HVFB sample, in 2012, 91.8 percent of the 2008 sample was re-interviewed (1,103 out of 1,163), yielding an annual attrition rate of 2.0 percent. These attrition rates are comparable to those in large-scale household surveys in developed countries.

We have examined whether household attrition is random or systematic. Table 2.2 provides an example of our approach. We estimate a probit where the dependent variable equals 1 if the household was interviewed in the 2012 survey round, and 0 otherwise. This is a function of baseline (2006) household characteristics and location (woreda) dummy variables.[4] In Table 2.2, parameter estimates have been converted into marginal effects, and standard errors account for clustering at the sampling (woreda) level. These results exclude the households in Amhara-HVFB.

A large number of household characteristics are not correlated with attriting. There is no association with the household head's age or schooling. Wealth, as measured by land and livestock holdings, does not affect the likelihood of attrition, nor does program participation or measures of households' connectedness to the area—the head being born in the locality and whether the parents of the head ever held an official position in this locality. Two characteristics do affect attrition. Female-headed households are two percentage points less likely to be traced, and all the household-size dummy variables are statistically significant. However, they are all measured relative to the omitted category, a one-person household. This tells us that one-person households are more likely to have attrited; since one-person households contain no children, this is unlikely to be a source of bias in our work. Relative to households of other sizes, attrition is less likely in households of six, seven, or eight members, but the magnitude of this difference—on the order of 1 to 3 percent—is unlikely to be meaningful.

Table 2.2 Determinants of the likelihood that a household was interviewed in the 2012 survey round

	Marginal effects	Standard error
Age of household head, years	0.000	0.000
Female household head, 0/1	-0.022**	0.010
Highest school grade completed by household head	-0.000	0.001
2 members in household, 0/1	0.062***	0.009
3 members in household, 0/1	0.063***	0.010
4 members in household, 0/1	0.075***	0.011
5 members in household, 0/1	0.082***	0.009
6 members in household, 0/1	0.093***	0.007
7 members in household, 0/1	0.086***	0.006
8 members in household, 0/1	0.082***	0.005
9 members in household, 0/1	0.073***	0.005
10 members in household, 0/1	0.067***	0.007
11 members in household, 0/1	0.069***	0.004
12 members in household, 0/1	0.056**	0.026
Landholdings operated (hectares)	0.001	0.007
Livestock owned, number	0.001	0.001
Household has been a PSNP beneficiary, 0/1	0.005	0.008
Household head was born in this locality, 0/1	0.012	0.016
Parent of head ever held an official position in this locality, 0/1	-0.010	0.021

Source: Authors' calculations.
Notes: PSNP = Productive Safety Net Programme. *Statistically significant at the 10% level; **statistically significant at the 5% level.

[4] For brevity, the parameter estimates for the woreda dummy variables are not reported in Table 2.2.

3. THE PRODUCTIVE SAFETY NET PROGRAMME

This section provides background information on the PSNP, drawing heavily on the extensive work we have done for the government of Ethiopia on the operations of the program (Berhane et al. 2011, 2013; Gilligan et al. 2007, 2009). After providing an overview of the program, we examine three aspects of program implementation relevant to our impact analysis: targeting, payments, and work. We then summarize some implications of this for our analytical work.

3.1. Program Overview

Between 1993 and 2004, the government of Ethiopia launched near-annual emergency appeals for food aid and other forms of emergency humanitarian assistance. While these succeeded in averting mass starvation, especially among the asset-less, they did not banish the threat of further famine and they did not prevent asset depletion by marginally poor households affected by adverse rainfall shocks. Further, the ad hoc nature of these responses meant that the provision of emergency assistance—often in the form of food-for-work programs—was not integrated into ongoing economic development activities. Starting in 2005, the government of Ethiopia and a consortium of donors implemented a new response to chronic food insecurity in rural Ethiopia. Rather than annual appeals for assistance and ad hoc distributions, a new program called the Productive Safety Net Programme (PSNP) was established. The program began at scale in the four regions covered by this study; there was no graduated scale-up or rollout.

The objective of the PSNP is “to provide transfers to the food-insecure population in chronically food-insecure woredas in a way that prevents asset depletion at the household level and creates assets at the community level” (Ethiopia, MoARD 2004, 2009a). Unlike the annual emergency appeals, the PSNP was conceived as a multiyear program so as to provide recipients with predictable and reliable transfers. Most beneficiary households perform public works; criteria for selection are that these households are poor (for example, they have low holdings of land or cattle) and food insecure, but that they also have able-bodied labor power. A much smaller proportion of beneficiaries receive direct support; these households are poorer than those receiving public works employment and lack labor power (this category includes households whose primary income earners are elderly or disabled). From 2005 through 2007, the public works component paid beneficiaries either 6 birr per day (increased to 8 birr in 2008, 10 birr in 2010, and 14 birr in 2012) in cash or 3 kilograms of cereals for work (depending on where they lived) on labor-intensive projects building community assets. Most activities occur between the months of January and June so as not to interfere with farming activities that occur in the second half of the year.

Initially, the PSNP was complemented by a series of food security activities, collectively referred to as the Other Food Security Programme (OFSP). While the PSNP is designed to protect existing assets and ensure a minimum level of food consumption, the OFSP was designed to encourage households to increase incomes generated from agricultural activities and to build up assets. The OFSP included access to credit; assistance in obtaining livestock, small stock, bees, tools, or seeds; and assistance with irrigation or water-harvesting schemes, soil conservation, or improvements in pastureland. However, relatively few households have had access to the OFSP. Given these problems, the Ethiopian government, in collaboration with donors, extensively redesigned the OFSP and christened the new program the Household Asset Building Programme (HABP). The HABP placed increased emphasis on contact and coordination with agricultural extension services while expanding access to credit through microfinance institutions and rural savings and credit cooperatives (Ethiopia, MoARD 2009b). This has led to an improvement in support provided by development agents. While many households reported contact with development agents, assistance remains concentrated on crop production. There is limited capacity to assist households with non-agricultural enterprises. Access to new forms of credit also has been limited. Relatively few households reported borrowing money to purchase inputs or to buy livestock (Berhane et al. 2013).

3.2. Targeting

The PSNP uses a mix of geographic and community-based targeting to identify chronically food-insecure households in chronically food-insecure woredas. When the program began, 190 woredas (roughly equivalent to a counties or districts) were selected on the basis of historical data on food aid allocations. These data were also used to determine the number of eligible beneficiaries in each region and woreda. This meant that coverage of the program was higher in the historically famine-prone northern regions of Tigray and Amhara and lower in the central region of Oromiya and the Southern Nations, Nationalities, and Peoples’ Region (SNNPR).

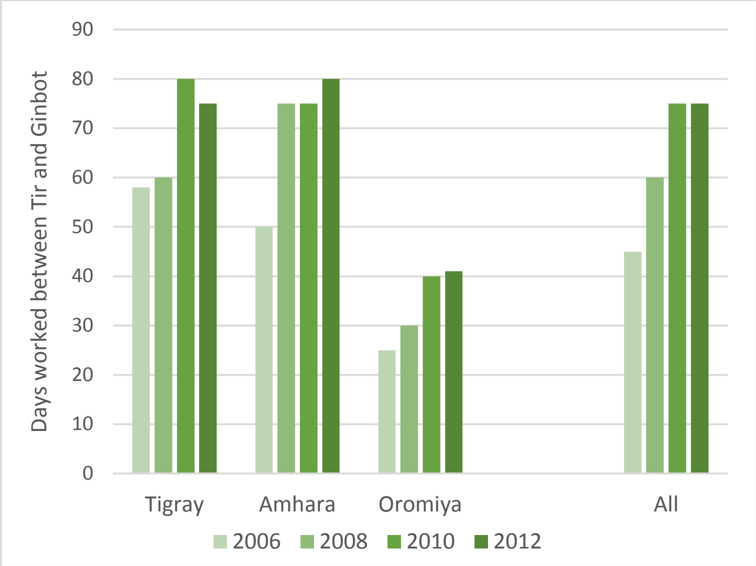
Initially, household-level targeting for the PSNP focused on selecting households that had high levels of food insecurity and that had been recipients of past emergency food aid. Having made the initial selection using the basic criteria, the program developers then verified and refined the selection of eligible households based on household assets (landholdings, oxen) and income from non-agricultural activities and alternative sources of employment. However, communities were given substantial discretion to modify this approach and to update their lists of food-insecure households annually based on local criteria. So, for example, households that suddenly become more food-insecure as a result of a severe loss of assets and are unable to support themselves, as well as any households without family support and other means of social protection and support, can be included in beneficiary lists. After PSNP eligibility is determined based on these criteria, households are assigned to public works or direct support: Eligible households with able-bodied adults receive transfers for their participation in public works projects, while those households that cannot provide labor or other means of support receive unconditional transfers. Most beneficiary households participate in public works; a much smaller proportion receive direct support.

Analysis of the PSNP's targeting indicates that it performs well. Coll-Black et al. (2012) demonstrate that the PSNP is targeted toward households that are both food-insecure and poor in terms of total household resources. Although there are some regional variations, overall program guidelines regarding targeting were followed. Public works projects targeted the poor for participation, rather than food-insecure households, but as poverty is highly correlated with food insecurity, food-insecure households were targeted as well. The program targeted direct support toward households with limited labor endowments, rather than targeting based on poverty. Over time, community understanding of targeting criteria improved across most PSNP regions. Households' identification of poverty-related factors as a reason why households are selected for public works improved in most regions, while it was increasingly well understood that the elderly and disabled are the intended recipients of direct support. Family or friendship connections were not reported as major factors in a household's likelihood to receive public works benefits or direct support.

3.3. Public Works Employment

Figure 3.1 displays the median number of days worked on PSNP public works per household, by region and year, conditional on at least one member of the household having worked at least one day in that year. It is important to note that because of the timing of the 2006, 2008, 2010, and 2012 surveys, these data pertain to the first five months of each year, the months of Tir, Yekatit, Megabit, Miaza, and Ginbot in the Ethiopian calendar (approximately January 9–June 9).

Figure 3.1 Median number of days worked per public works beneficiary household between the months of Tir and Ginbot (early-January to early-June), by region and survey year



Source: Authors' calculations.

Figure 3.1 indicates that there are some regional and temporal variations in days worked over this period. Employment was lowest when the program began in 2006 and rose in 2008 and 2010; by 2010, the median beneficiary household received 75 days' employment via the PSNP. This level, 75 days, is also recorded in 2012. In all years, employment in this period is higher in Tigray and Amhara than in Oromiya. In surveys fielded after 2006, respondents were asked about employment in the previous year (so in the 2008 survey, respondents were asked about work in 2007; in the 2010 survey, respondents were asked about work in 2009, and so on). This allows us to assess whether, for example, employment levels in Oromiya are low or, instead, whether employment occurs slightly later in the calendar year. Analysis of these data showed that in Tigray and Amhara, approximately 75–80 percent of PSNP employment was completed between Tir and Ginbot, compared to 40 percent in Oromiya.

In preliminary work, we examined the distribution of employment by sex between Tir and Ginbot. In Tigray, employment is fairly evenly split between men and women. Around 40 percent of public works in Amhara is undertaken by women. Women's participation in public works is lowest in Oromiya, where only about 30 percent of women worked.

Finally, we examined the age distribution of public works participants; in particular, we wondered if children were employed in public works activities. This does not appear to have been the case. For example, the quantitative data show that there are only 10 children among the more than 2,500 individuals who undertook public works employment in the first five months of 2010. Individuals aged 20 to 49 years make up nearly 70 percent of the workforce engaged in public works.

3.4. Payments

In the first five years of the program, there were difficulties in making timely and predictable payments to beneficiaries. Since 2010, however, program performance has significantly improved. A full discussion of this issue can be found in Berhane et al. (2011, 2013) and Gilligan et al. (2007, 2009).

Data on payments were collected in all survey rounds. Respondents who indicated that they were PSNP participants were asked to recall, month by month, the payments they had received. Payments received in kind (mostly grain, but also some pulses and oil) were valued using local market prices obtained from a market survey fielded alongside the household survey. Relying on respondent memories of past payments raises the issue of measurement error in these data. We address this, albeit partially, by taking advantage of the fact that by the time we fielded the 2012 survey, the PSNP was rolling out "client cards" that recorded household payments. Enumerators were instructed to ask to see the respondents' client cards and copy information from those onto the questionnaire. If the beneficiaries stated that they did not have a client card, or that they did have a client card but were unable or unwilling to show it, enumerators were instructed to ask the respondents to recall their payment information.

Using these data, we reproduce a table from Berhane et al. (2013). Table 3.1 compares the level and distribution of public works payments data from these two sources for 2011. For public works, mean payments were 2,387 birr for 2011, when this information was collected from the client card, and 2,342 birr, when obtained through recall. This is a difference of only 45 birr, or 1.9 percent. Further, the two sources of information provide remarkably similar distributions of payments and similar means and distributions when we disaggregate by recall period. For direct support payments, the difference in mean values is only 10 birr, or 0.9 percent (full results available on request). This gives us further confidence that our payments data are providing a reasonable representation of payments actually received by beneficiaries.

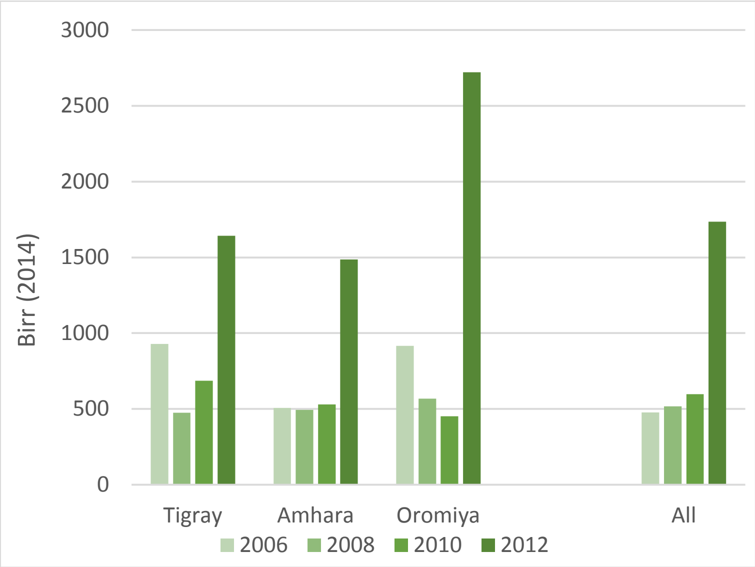
Table 3.1 Comparison of public works payment data from client cards and respondent recall, by time period and type of transfer, 2011

Source	Mean	Sample size	Percentiles p10	p25	p50	p75	p90
Tir–Tahisis (January–December) 2011							
Recall	2,342	878	790	1,225	1,956	2,940	4,250
Client card	2,387	537	840	1,260	1,845	3,000	4,200
All	2,359	1,415	825	1,260	1,917	2,977	4,200
Tir–Sene (January–June) 2011							
Recall	1,852	868	600	975	1,500	2,400	3,364
Client card	1,816	533	600	981	1,500	2,160	3,350
All	1,838	1,401	600	981	1,500	2,300	3,353
Hamle–Tahisis (July–December) 2011							
Recall	1,045	429	250	444	825	1,363	1,960
Client card	1,078	291	240	560	900	1,350	1,800
All	1,059	720	245	471	883	1,350	1,825

Source: Household questionnaire 2012.

Median payments to PSNP public works participants by survey round are shown in Figure 3.2. Because there was significant food price inflation during the period under study, we constructed a price index using information on grain prices collected as part of the quantitative community questionnaire. We used this to deflate all payment values to 2014 birr. This shows that in real terms payments to beneficiary households were largely flat between 2006 and 2008, even though the number of days worked was rising (Figure 3.1). Payments rose somewhat more between 2008 and 2010 (by 80 birr), but the major rise occurred between 2010 and 2012.

Figure 3.2 Median total payments to public works beneficiary households, by region and survey year



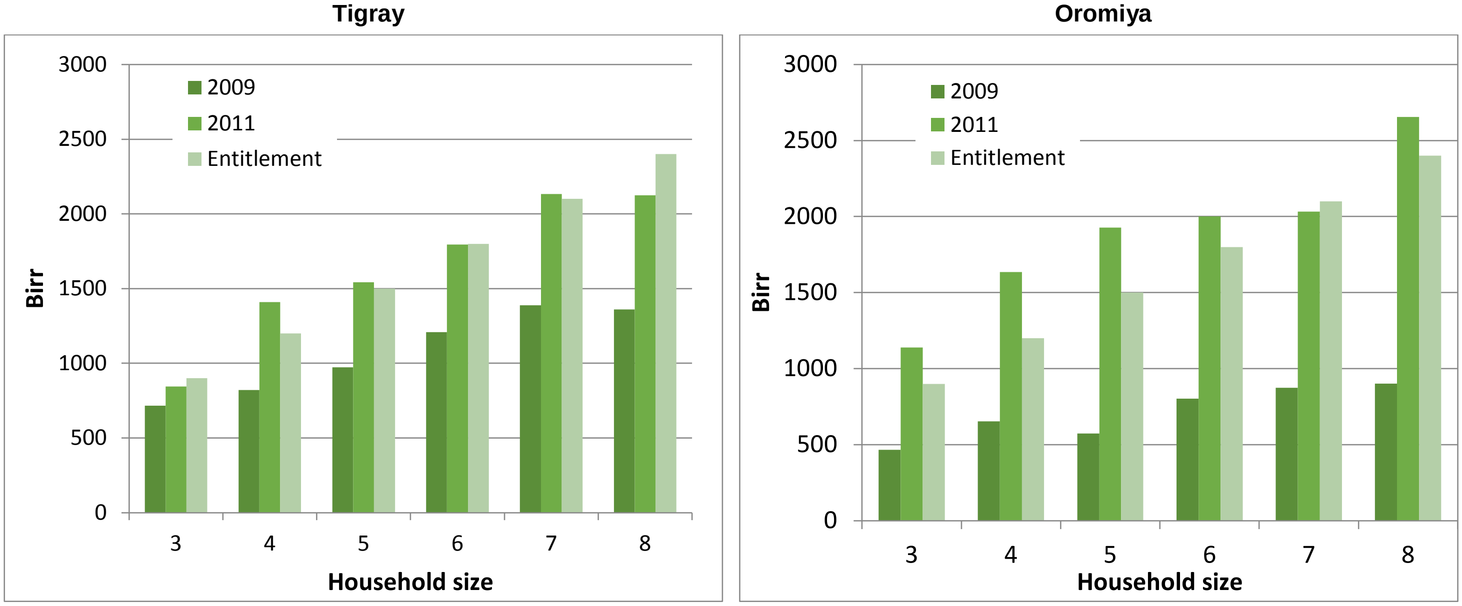
Source: Authors' calculations.

Why do we observe this pattern? In the early years of the program, there were significant delays in making payments. For example, Gilligan et al. (2009) show that just over 80 percent of participants in public works in the SNNPR received at least 80 percent of the money owed to them in the first five months of 2007. The comparable figures for Oromiya, Amhara, and Tigray are 63 percent, 62 percent, and 18 percent, respectively. Generally, the arrears were not subsequently paid. Other problems included difficulties in processing payment information, difficulties in ensuring timely receipt of food needed for payments made in kind, and limited access to the transport needed to pay PSNP beneficiaries.[5]

In response to these problems, the government of Ethiopia revised the payment system and introduced the concept of "full family targeting" (FFT). Under FFT, each member is entitled to five days' work per month for six months. For example, at a wage rate of 10 birr per day, this yields a payment of 300 birr per household member. This means that a three-person household should receive 900 birr, a four-person household should receive 1,200 birr, and so on. In our operational work for the government of Ethiopia, we compared these expected levels of payment against mean total payments by region for households ranging in size from three to eight persons. Figure 3.3 shows examples of this work for two regions, Tigray and Oromiya. These figures show considerable variation in payments by region and year. They also show considerable improvement in payments between 2010 and 2012.

[5] We have limited information on why some localities received cash payments while others received food. Regions and woredas could request either cash, food, or a mixture. However, the actual decision to provide either or both forms of payment depended on the physical availability of commodities in government storage facilities, which varied over time and space. Requests for in-kind transfers increased when food prices rose in 2008 but, we understand, subsided thereafter. Further, the timing, frequency, and amount of in-kind transfers differed from cash payments. Since we cannot model the determinants of the receipt of cash or in-kind payments, and because they are not directly comparable, we do not attempt to disaggregate our results by form of payment.

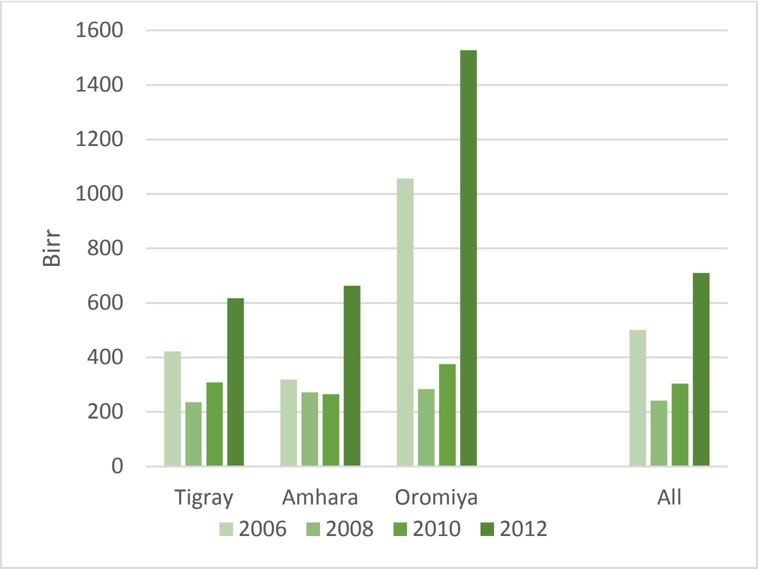
Figure 3.3 Comparison of normalized total payment to full family targeting entitlement, by household size, Tigray and Oromiya, 2009 and 2011



Source: Berhane et al. (2013).

Finally, we consider payments made to households receiving direct support. These are shown in Figure 3.4. While they show the same pattern over time as those for public works beneficiaries, these payments are considerably lower than those found in Figure 3.2.

Figure 3.4 Median total payments to households receiving direct support, by region and survey year



Source: Authors' calculations.

3.5. Linking the PSNP to Children's Nutritional Status

Mindful of this description of the implementation of the PSNP, we can link the program to standard economic models of the determinants of children's nutritional status. We conceptualize parental decisions to devote resources to improving child nutrition as being motivated both by immediate concern about the welfare of their children and by longer-run concerns about investing in the human capital of their children. Household welfare is assumed to increase as children's nutritional status improves, although possibly at a diminishing rate. Moreover, increases in certain measures of nutritional status, such as body mass, may be associated with reductions in welfare beyond a certain point. Parents may not have identical preferences regarding the use of family resources, but could engage in (perhaps implicit) bargaining about such allocations, in which the strength of the bargaining position of each parent may depend on her or his access to resources, including those provided by social networks and policies.

Decisions that parents make, whether through bargaining or some other mechanism, about devoting resources to their children's nutrition and health, are constrained in several ways. There are resource constraints reflecting income and time available as well as prices faced by households. There is also a constraint arising from the production process for health outcomes, including nutritional status. This constraint links nutrient intakes—the physical consumption of macronutrients (calories and protein) and micronutrients (minerals and vitamins); time devoted to the production of health and nutrition; locality characteristics, such as the presence of preventive and curative health facilities and the prevalence of infectious diseases; the individual's genetic makeup; and knowledge and skill regarding the combination of these inputs to produce nutritional status.

Participation in the public works component of the PSNP affects both the income and time constraint but in different ways. The income received from undertaking this work relaxes the income constraint. To the extent that some of this income is used to improve the consumption of nutrients, this should improve the nutritional status of preschool children. However, participation in public works creates a new demand on parents' time. If this increased workload reduces the time devoted to childcare (for example, if the frequency of child feeding is reduced because mothers are away from the household), then this aspect of program participation may harm preschool nutritional status of children.

3.6. Summary

This section has provided a short description of the design and operation of the PSNP, drawing heavily on our operational work undertaken for the government of Ethiopia. Two aspects of program implementation are relevant for our analysis. First, the program was well targeted, with the result that beneficiaries are, in fact, poorer than non-beneficiaries. This means that we should be especially cautious when comparing unconditional means across PSNP participants and non-participants. Second, we will need to be careful about exactly how we define program participation—being a PSNP household in 2008 was clearly very different in terms of work requirements and payments compared to being a PSNP household in 2012, and benefit levels also differ markedly between PSNP public works and PSNP direct support households.

4. IMPACT EVALUATION ISSUES

4.1. Overview

The fundamental problem for a quantitative impact evaluation of a program like the PSNP is that we observe only what happens to beneficiaries who are receiving benefits; we do not observe what would happen to the same households if they did not receive benefits. This is called the problem of the counterfactual. A second issue is selection bias. Selection bias arises when beneficiaries differ in some systematic way from non-beneficiaries. Our ability to make statements about the causal impact of the PSNP rests on how well we can address these two problems. Broadly speaking, there are four methods for doing so:

- A randomized evaluation design
- A regression discontinuity design
- Instrumental variables
- Matching estimators

We consider these in turn.

Randomized designs are attractive because by design, under certain conditions, they resolve the selection issue. However, from the very outset, the government of Ethiopia refused to consider a randomized design for the evaluation of the PSNP.

A regression discontinuity design (RDD) requires a cut-off or threshold that separates beneficiaries from non-beneficiaries. The means by which the PSNP is targeted—the use of general criteria interpreted by local authorities (see Section 3)—precludes the use of RDD. In our evaluation work, both quantitative and qualitative, we explored with local officials how they implemented targeting in the PSNP. These discussions made clear that there was no simple quantitative mechanism used to select beneficiaries and hence no cut-off or threshold. RDD was not feasible.

As is well known, instrumental variables (IV) require the identification of variables that affect the treatment variable directly but do not affect the outcome variable. IV estimation is typically implemented in stages by running two regressions. In the first stage, we regress the treatment variable on the instrumental variables and variables that capture basic household and child characteristics. This first-stage regression gives us the prediction of the treatment based on the instruments and the basic household and child characteristics. In the second stage, we regress the outcome variable of interest on the *prediction* of the treatment variable obtained from the first-stage regression and the basic household and child characteristics.

The IV approach is valid if the instruments satisfy two assumptions: (1) the instruments are good predictors of the treatment variable (for example, the level of payments) and (2) the instruments correlate only with the treatment variable and not with the outcome variable of interest other than through the treatment variable. We experimented extensively with IV estimators but could not find instruments that had sufficient explanatory power to predict the first-stage outcomes.

These considerations leave us with matching methods.

4.2. Matching Methods, Including Inverse-Probability-Weighted Regression-Adjustment Estimators

Matching methods of program evaluation construct a comparison group by “matching” treatment households to comparison group households based on observable characteristics. The impact of the program is then estimated as the average difference between the outcome for each treatment household and a weighted average of outcomes in each similar comparison group of households from the matched sample.

When we began PSNP work in 2006, we were aware that other impact evaluation methods were not likely to succeed, so we ensured that the surveys we implemented could support the use of matching methods. Matching methods provide reliable, low-bias estimates of program impact under the following conditions: (1) the same data source is used for participants and non-participants, (2) the data include meaningful X variables capable of identifying program participation and outcomes, and (3) participants and non-participants have access to the same markets (Heckman, Ichimura, and Todd 1997, 1998). Since the same survey instrument was applied everywhere, criterion (1) was satisfied. The biannual PSNP surveys were designed to include a rich set of variables that will identify program participation and outcomes related to child nutrition and other outcomes of interest as required by criterion (2). Criterion (3) was met by sampling treatment and control households within the sample kebeles, as noted in Section 2.

Matching can be accomplished in several ways. The best-known method is propensity score matching (PSM). PSM uses a fully specified probit regression to estimate the treatment model, or the process by which respondents are selected into the treatment or comparison groups. It then compares each treatment observation to only one control observation in computing the individual treatment effect. PSM uses a fully nonparametric technique to estimate the outcome model. The individual treatment effect is calculated as a simple difference between the outcome for the treatment unit and its nearest control unit; this estimate does not control for other variables that may also affect the outcome variable.[6]

In preliminary work, we estimated program impacts using PSM. However, the results reported in this document are based on a matching method that improves on PSM: inverse-probability-weighted regression adjustment (IPWRA) (e.g., Imbens and Wooldridge 2009). IPWRA improves on PSM in two ways. First, the outcome model in IPWRA is fully specified and can include controls for the observation’s concurrent or baseline characteristics. For example, IPWRA allows the researcher to explicitly control for whether the child in the outcome model is male or female; because PSM looks only at the difference between each treated unit and its nearest control unit as measured by the propensity score, it does not explicitly control for the child’s gender unless the child’s gender is included in the treatment model. The improvement in efficiency due to the inclusion of these control variables in IPWRA over PSM is analogous to the improvement in precision one finds when including additional covariates in the evaluation of a randomized control trial. While comparing the difference between outcomes in the randomly selected treatment and control groups is unbiased, including covariates in addition to the treatment status absorbs variance and thus allows a more precise estimate of the treatment effect. A further benefit is that it

[6] A second method is nearest neighbor matching (NNM) (see Abadie and Imbens 2006). Differences between NNM and PSM derive primarily from the rule used to select comparable non-beneficiaries and the weights used to construct the difference in weighted average outcomes. NNM, a form of “covariate matching,” matches beneficiaries to non-beneficiaries based directly on observable characteristics. Each beneficiary is matched to the group of non-beneficiaries with the smallest average difference in preprogram characteristics, where this difference is determined using a multidimensional metric across all control variables.

is no longer necessary to ensure balance across the baseline covariates that appear in the probit used to estimate the propensity scores, as these also appear in IPWRA. Second, PSM compares each treatment observation to only one or a few control observations that have a similar likelihood of being treated. In essence, PSM puts a weight of 1 on the nearest control observation and a weight of 0 on all other observations. IPWRA implicitly compares every unit to every other, while placing higher weights on observations that have a similar likelihood of being in the treatment or comparison group and lower weights on observations that are dissimilar. Because more observations are included in the model that compares a treatment unit to its hypothetical counterfactual, statistical precision is increased.

In addition to gains in efficiency, IPWRA has one additional attractive feature in comparison to PSM: it is doubly robust. Consider first PSM. If the treatment model is mis-specified (that is, the model is missing a variable or the functional form is incorrect), PSM will provide inconsistent estimates. With IPWRA, by contrast, if the treatment model is mis-specified, the estimates of the treatment effect will still be consistent so long as the outcome model is not also mis-specified. The reverse is also true: if the treatment model is appropriately specified but the outcome model is mis-specified, IPWRA still delivers consistent estimates. While we are confident in all our specifications, we appreciate this double-robust property as a fallback (Imbens and Wooldridge 2009).

IPWRA is accomplished in three steps. First, the probability that an observation is treated is estimated using a treatment model, usually with a probit or logit regression. The predicted probabilities are used to reweight the sample by the inverse of the probability that each observation is in the treatment or control group. Second, the expected outcome is estimated for each observation using a weighted outcome model that includes both the observable characteristics used to estimate the treatment model and additional information. For example, if the outcome of interest is a child's height-for-age z-score, the outcome model may include the child's age in addition to the household demographic characteristics that were included in the treatment model. Baseline data on outcomes can also be used in this way to more precisely estimate treatment effects at endline. The outcome model is used to predict the expected outcome for each observation twice: once from the perspective (weights) of the probability of being treated and again from the perspective (weights) of the probability of being in the control group. Finally, the average outcome for treatment and control observations is calculated. The difference between these two averages is the estimated treatment effect.

To see how IPWRA works, consider a very simple model of child height (Y), where this outcome is a function of child age (W). We have two groups of households, PSNP beneficiaries (B = 1) and non-beneficiaries (B = 0). We estimate these models of height for the two groups separately.

$$Y_{B=1} = \alpha_{B=1} + \beta_{B=1} W + \varepsilon_{B=1} \qquad (1)$$

and

$$Y_{B=0} = \alpha_{B=0} + \beta_{B=0} W + \varepsilon_{B=0} \qquad (2).$$

We could estimate (1) and (2) separately and calculate predicted values for $Y_{B=0}$ and $Y_{B=1}$. Having done so, it would be tempting to take the difference in these predicted values and call that the impact of the PSNP. The problem, of course, is that beneficiaries are not randomly selected; there is correlation, for example, between $\varepsilon_{B=1}$ and $\alpha_{B=1}$. This can be resolved by weighting these regressions (see Imbens and Wooldridge [2009, 38–39], where the weights are derived from the inverse propensity scores). This yields the average treatment effects on the treated (ATET).

4.3. Implementing the Inverse-Probability-Weighted Regression-Adjustment Estimator

Implementing IPWRA requires that we precisely define what is meant by a PSNP beneficiary and that we estimate a model that predicts program participation. To do so, we draw on our earlier discussion of the PSNP.

To begin, recall that the PSNP has two components: public works and direct support. These differ in a number of important ways. Public works has a work requirement, while direct support is an unconditional transfer. The criteria for selection into these components are somewhat different. Public works is for households that are poor (for example, they have low holdings of land or cattle) and food-insecure but that also have able-bodied labor power. Direct support households are poorer than those receiving public works employment and lack labor power; they include those whose primary income earners are elderly or disabled. Direct support households received much lower payments than public works beneficiaries.

There is one additional consideration. Given the targeting used to identify direct support households, they tend to have few children—in fact, more than 85 percent of the children in our sample are found in public works households. So a question that arises is whether (1) we include direct support households as part of the treatment group (which would mean that we are assessing the impact of any type of household participation in the PSNP relative to nonparticipation); (2) we define treatment as participation in the public works component and include all other households (including direct support) as potential controls; or (3) we define treatment as participation in the public works component, exclude all direct support households, and use households that receive no PSNP benefits as controls. In preliminary work, we experimented with all three possibilities. Qualitatively (in terms of signs, magnitudes, and statistical significance) it makes no difference which of these we use—an unsurprising result, given that direct support households have such few children. In the results below, we use (2). We define beneficiaries as households that participate in the public works component of the PSNP. Our comparison households are households that are not public works beneficiaries.

The next issue we need to confront is whether we should pool our data across all survey rounds or estimate impacts by year. Payments data shown earlier indicate that there are significant differences in program implementation across years, making pooling of data inadvisable. For this reason, we assess impact by year.

We use a probit model to predict program participation. (In preliminary work, we also used a logit model; our results are not affected by the choice of prediction estimator.) Based on the targeting criteria for the PSNP, along with our assessment of how these criteria have been implemented (see Section 3), the following covariates are used as predictors: household livestock holdings 12 months prior to the survey; household landholdings; age, sex, and educational attainment of the household head; the number of males aged 16 to 60 years resident in the household; and the number of females aged 16 to 60 years resident in the household. Schooling of the household head, livestock holdings, and landholdings capture dimensions of household wealth, and these variables are strongly correlated with measures of household food security. We use livestock holdings as of 12 months prior to the survey to avoid the possibility that these are affected by receipt of PSNP transfers. (Note that in Ethiopia, land is owned by the state and is allocated to households on a usufruct basis and so holdings are not affected by PSNP status.) Age affects access to public works, as older households are more likely to receive direct support payments. Finally, since the public works component requires that households have labor power, we include as proxies for this, the number of adult males and adult females aged 16 to 60 years.

Table 4.1 reports the results of this probit for the survey round that was fielded in 2012. (Results for earlier rounds are available on request.) In addition to the covariates described above, we include (but do not report) locality (woreda or district) level dummy variables. This allows us to condition out local factors—such as the budget allocated to the PSNP in different localities—in the estimation of the predicted probabilities. Parameter estimates have been converted into marginal effects. Standard errors are clustered at the level of the sampling unit.

Table 4.1 Probit estimates (marginal effects) of correlates of participation in Productive Safety Net Programme (PSNP) public works, 2012

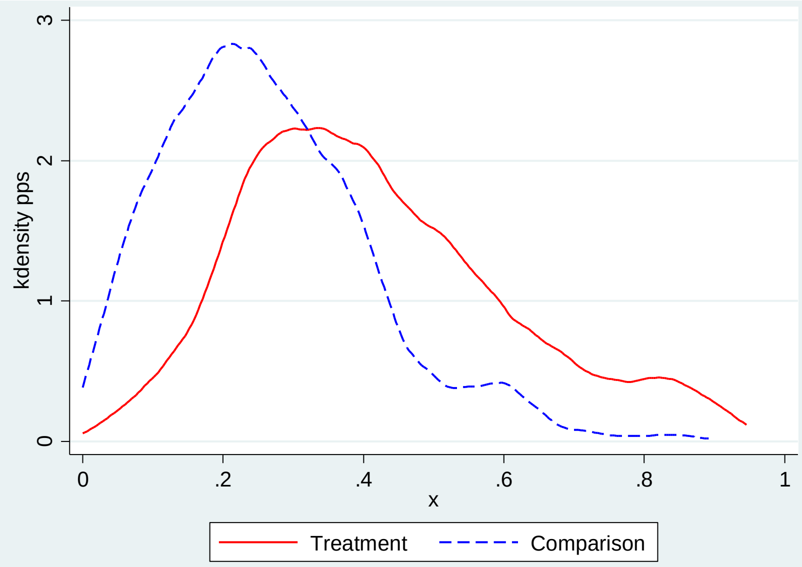
Explanatory variable	Marginal effects	Standard error
Livestock holdings, 2011, Tropical Livestock Units	-0.017***	(0.007)
Land operated by the household, 2012, ha	-0.024***	(0.009)
Highest grade attained, household head, 2012	-0.011**	(0.006)
Age of household head, 2012, years	-0.004***	(0.001)
Female household head, 2012, 0/1	0.063*	(0.036)
Males in household aged 16–60, 2012	0.016**	(0.008)
Females in household aged 16–60, 2012	0.025***	(0.009)

Source: Authors' calculations.
Notes: *Significant at the 10% level; **significant at the 5% level. Standard errors clustered at the kebele level. Woreda dummy variables included but not reported.

Table 4.1 shows that the likelihood of participation falls for wealthier households (the coefficients on livestock, land, and household head's schooling are negative and statistically significant) and rises with the availability of labor power (number of adult males and females).

A requirement for the use of inverse propensity scores is that there is common support. That is, the probability of being a participant (nonparticipant) is both nonzero and less than one for all observations. One way of assessing this is to plot the propensity scores for both participants and non-participants and see if the distributions of these overlap. Figure 4.1 shows that these distributions do, in fact, overlap.

Figure 4.1 Density functions showing common support, public works and comparison households, 2012



Source: Authors' calculations.

Before moving to our results, we note the following. As with any method of estimating treatment effects, several assumptions are needed to justify the use of IPWRA. First, the conditional independence assumption must hold for the estimation of average treatment effects. This assumption states that no unobservable variable affects both the likelihood of treatment and the outcome of interest, after conditioning on covariates. Because IPWRA includes more covariates (in the outcome model) than does PSM (which includes only the covariates in the treatment model), this assumption is more likely to hold with IPWRA than with PSM. Second, the i.i.d. ("independent and identically distributed" observations) assumption must hold. This assumption means that the potential outcomes and treatment status of each individual are independent of the potential outcomes and treatment status of all other individuals in the sample. This assumption also must hold for both IPWRA and PSM. Third, the overlap assumption must hold. This assumption states that every observation in the sample must have a positive estimated probability of being treated. This assumption must hold for both IPWRA and PSM, and because the treatment models of IPWRA and PSM are often estimated using the same method (for example, probit or logit models), the assumption is theoretically equivalent for both methods. Note that the statistical package we employ for our work (Stata 14) automatically restricts observations to those in the common support. The plots of the estimated likelihood of treatment shown in Figure 4.1 indicate that we have common support.

5. THE IMPACT OF THE PSNP ON CHILDREN'S NUTRITIONAL STATUS: 2008–2012

5.1. Data

In the 2008, 2010, and 2012 survey rounds, anthropometric measures were obtained for all children living in the household who were aged 6 months to 5 years. Enumerators were instructed to measure lengths for children younger than 24 months and heights for children 24 to 60 months. Weights were obtained with children wearing light clothing. Data from all rounds and regions have been cleaned. Table 5.1 gives the number of children measured by round and sex.

Table 5.1 Number of children whose height and weight were measured, by round and sex

Number	2008	2010	2012
Girls	1,722	1,556	1,317
Boys	1,677	1,561	1,355
All	3,399	3,117	2,672

Source: Authors' calculations.

Measurements of height and weight were converted to z-scores using the World Health Organization (WHO) growth standards (WHO 2006; de Onis et al. 2007). These standards allow us to assess child height and weight relative to well-nourished children of the same age and sex. A z-score expresses these measures in terms of standard deviations. For example, the height-for-age z-score is calculated by taking the child's height and subtracting the median height of a reference population of children of the same age and sex and dividing this by the standard deviation for that reference population. Children with positive z-scores have heights above this reference population. Children with negative z-scores have heights below this reference population; see de Onis et al. (2007) for details.

Children are considered to be chronically undernourished (stunted) if they have a height-for-age z-score below –2, that is, their height given their age and sex is two standard deviations below the median height for a child of the same age and sex. Chronic undernutrition reflects the malign synergistic effects of continued inadequate food intake together with repeated infection. Over a protracted period of time, the child's body fails to receive sufficient nutrients—calories and micronutrients—to grow or the need to fight repeated infections diverts energy that otherwise would be used for child growth. Physical growth lost in the first two years of life is never fully recovered. There is a growing body of evidence indicating that chronic undernutrition leads to irreversible neurological damage that adversely affects schooling and economic productivity (Hoddinott et al. 2013). Children are considered to be acutely undernourished (wasted) if they have a weight-for-height z-score below –2, that is, their weight given their height is two standard deviations below the median height for a child of the same age and sex. Acute undernourishment is more sensitive to recent episodes of illness or inadequate food intake.

Several features of these measurement data are worth bearing in mind. Instructions on how to weigh children were widely followed, with 91 percent of children weighed wearing light clothing and no meaningful differences in this practice across years. The practice of measuring lengths rather than heights for children under 24 months was largely followed for very young children, but enumerators did not always strictly adhere to these guidelines with older children. For example, approximately 95 percent of 6-month-old children were measured recumbently, compared to 50 percent of 23-month-old children. Across all children, 74 percent of children in the 6- to 23-month age group were measured lying down. This variation has the potential to introduce some measurement error into the height data. A much greater concern, however, is that few births are registered—only 1.5 percent of children had a birth certificate and only 3.8 percent had the child's birth date recorded on a clinic card. In the vast majority of cases, 86 percent, caregivers gave their best recollection of when the child was born. Few could give exact information—date, month, and year. Since we describe children's heights relative to their age and sex, this introduces measurement error into these data. This problem is compounded by the tendency of mothers to report ages as full years, something particularly pronounced in older children. In our analysis, we address the potential confounding effects of these measurement errors in several ways. First, we follow the convention of dropping children with z-scores below –5.5 and above 5.5. Second, we assess the robustness of our empirical results by controlling for how data on ages were obtained.

In the 2010 and 2012 survey rounds, data were also collected on the diets of children 6–24 months old. The format was similar to that used in other nutrition surveys in Ethiopia. Enumerators had a list of foods (for example, eggs) or food groups (for example, any foods made from beans, peas, lentils, nuts, or seeds). For each item, mothers were asked if their children had consumed that food yesterday during the day or at night.

The 2010 and 2012 surveys also included a brief section asking if a child 6–24 months old had experienced any of the following symptoms in the previous two weeks: fever, cough or cold, fast breathing or shortness of breath, or diarrhea.

5.2. Basic Descriptive Statistics

Table 5.2 shows the mean values for children's height-for-age z-score (HAZ), stunting, weight-for-height z-score (WHZ), and wasting by survey round. There was no change in mean HAZ between 2008 and 2010, although between 2010 and 2012 it did improve by approximately 0.1 standard deviation (SD).[7] The percentage of stunted children rose slightly between 2008 and 2010 before falling in 2012; the difference in stunting prevalence between 2010 and 2012 is statistically significant at the 10 percent level. Mean WHZ declined by approximately 0.1 SD between 2008 and 2010. The percentage of children wasted drifted lower between 2008 and 2012, from 16.6 percent in 2008 to 15.5 percent in 2012. These values are broadly comparable to those found in the 2011 Ethiopia Demographic and Health Survey (DHS) (CSA and ICF International 2012),

[7] As discussed in section 2, there were some changes to the sample between 2008 and 2010 and again between 2010 and 2012. We reran these descriptive statistics, restricting the sample to children whose households were surveyed in all rounds. Doing so does not substantively change the results reported here.

which showed that in all rural areas, mean HAZ and stunting were –1.8 and 46.2 percent, respectively. However, while the 2011 DHS showed somewhat higher mean values of WHZ across all rural areas, –0.6, it also showed a lower prevalence of wasting, 10.2 percent.

Table 5.2 Mean values of height-for-age z-scores, stunting, weight-for-height z-scores, and wasting, by survey round

	Mean height-for-age z-score (HAZ)	Percent stunted	Sample size	Mean weight-for-height z-score (WHZ)	Percent wasted	Sample size
2008	-1.91	47.9	3,088	-0.53	16.6	3,203
2010	-1.90	50.9	2,893	-0.44	16.0	2,999
2012	-1.81	48.8	2,524	-0.43	15.5	2,580

Source: Authors' calculations.

In Table 5.3, we disaggregate these results by child sex. We observe similar temporal patterns for boys and girls in terms of our measures of chronic undernutrition, HAZ and stunting. Acute undernutrition measures, WHZ and wasting, improve for girls but not for boys. The 2011 Ethiopian DHS mean values for stunting are 42.5 percent for all girls (that is, rural and urban areas combined) and 46.2 percent for all boys. These are lower than our results, but this difference is to be expected, given that chronic undernutrition is lower in urban parts of Ethiopia.[8] Similarly, DHS values for wasting are also lower: 8.2 percent for girls and 11.1 percent for boys. Consistent with the DHS, in the PSNP data, unconditional mean values for chronic and acute undernutrition are worse for boys than for girls.

Table 5.3 Mean values of height-for-age z-scores, stunting, weight-for-height z-scores, and wasting, by survey round and sex

	Mean height-for-age z-score (HAZ)	Percent stunted	Sample size	Mean weight-for-height z-score (WHZ)	Percent wasted	Sample size
Girls						
2008	-1.87	49.6	1,529	-0.49	15.3	1,591
2010	-1.87	51.5	1,459	-0.40	14.2	1,505
2012	-1.78	48.5	1,291	-0.35	13.2	1,327
Boys						
2008	-1.94	52.5	1,559	-0.57	17.8	1,612
2010	-1.93	50.6	1,434	-0.48	17.7	1,494
2012	-1.85	49.3	1,233	-0.51	18.0	1,253

Source: Authors' calculations.

Table 5.4 disaggregates these descriptive data by survey round and region. Consistent with what we observe in the 2011 Ethiopian DHS, stunting is highest in Tigray and lowest in Oromiya. There are regional differences in wasting, but these are less pronounced than the differences in stunting. In Tigray, Amhara, Amhara-HVFB, and Oromiya, both HAZ and stunting worsen between 2008 and 2010 before improving between 2010 and 2012. The temporal pattern is slightly different in SNNPR, where chronic undernutrition improves between 2008 and 2010, but does not change between 2010 and 2012.

[8] The DHS does not report gender-disaggregated results for rural areas.

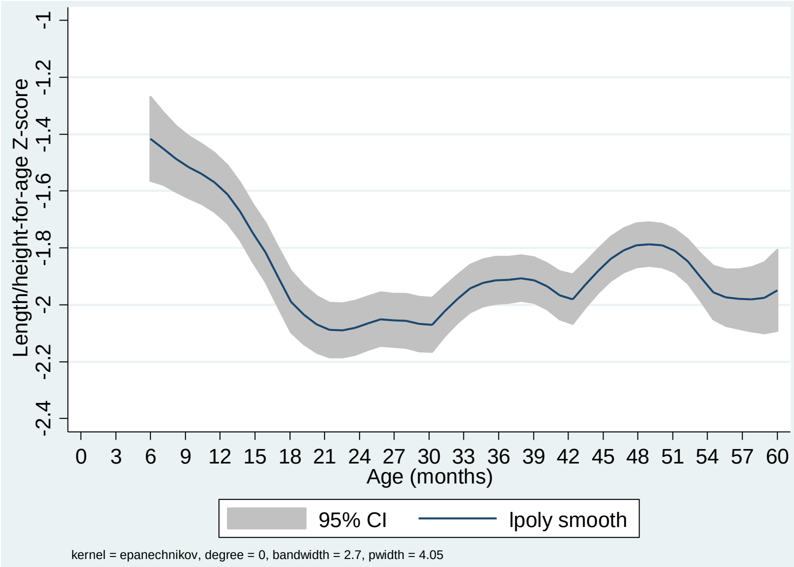
Table 5.4 Mean values of height-for-age z-scores, stunting, weight-for-height z-scores, and wasting, by survey round and region

	Height-for-age z-score			Stunting (%)		
	2008	**2010**	**2012**	**2008**	**2010**	**2012**
Tigray	-2.18	-2.29	-2.17	58.6	60.7	57.8
Sample size	*546*	*483*	*396*			
Amhara	-1.79	-1.86	-1.92	47.5	50.3	49.2
Sample size	*415*	*370*	*364*			
Amhara HVFB	-1.94	-2.05	-1.92	53.6	55.0	51.9
Sample size	*689*	*562*	*459*			
Oromiya	-1.71	-1.80	-1.53	46.1	47.9	42.8
Sample size	*750*	*770*	*682*			
SNNP	-1.94	-1.69	-1.76	50.2	46.2	47.3
Sample size	*641*	*660*	*581*			
Total	-1.90	-1.91	-1.81	51.1	51.4	48.9
Sample size	*3,041*	*2,845*	*2,482*			
	Weight-for-height z-score			**Wasting (%)**		
Tigray	-0.63	-0.57	-0.23	17.8	17.9	16.3
Sample size	*568*	*496*	*404*			
Amhara	-0.47	-0.50	-0.50	15.6	15.9	13.4
Sample size	*423*	*395*	*367*			
Amhara HVFB	-0.81	-0.63	-0.73	17.9	21.0	19.8
Sample size	*698*	*581*	*474*			
Oromiya	-0.47	-0.41	-0.63	16.6	15.9	17.4
Sample size	*783*	*804*	*691*			
SNNP	-0.27	-0.14	-0.06	14.8	9.7	11.1
Sample size	*683*	*669*	*602*			
Total	-0.53	-0.43	-0.43	16.6	15.9	15.6
Sample size	*3,155*	*2,945*	*2,538*			

Source: Authors' calculations.

We consider the pattern of children's nutritional status by child age. We use a nonparametric technique to smooth out these data; this also allows us to graph confidence intervals. Results are shown in Figure 5.1.

Figure 5.1 Height-for-age z-scores, by age



Source: Authors' calculations.

Figure 5.1 starts with children aged 6 months. Even at this early age, there is evidence of significant problems of chronic undernutrition, with the predicted mean HAZ at age 6 months already below –1.4. HAZ declines rapidly from ages 6 to 24 months, after which it essentially levels out, sometimes bouncing a little higher, sometimes a little lower, from 24 to 60 months. Ignoring these somewhat random fluctuations, what stands out is the swift decline from ages 6 to 24 months and just how bad chronic undernutrition is in this population.

Finally, we consider the unconditional mean estimates of children's nutritional status by survey round and PSNP status. Table 5.5 indicates that chronic undernutrition, using either HAZ or stunting as the measure, is slightly higher in PSNP households than non-PSNP households. This is true in all survey rounds. Acute undernutrition, using either WHZ or wasting, is essentially the same in PSNP and non-PSNP households.

Table 5.5 Height-for-age z-scores, stunting, weight-for-height z-scores, and wasting, by survey round and PSNP status

	Height-for-age z-score	Percent stunted	Weight-for-height z-score	Percent wasted
2008				
PSNP households	-1.87	49.7	-0.54	17.4
Non-PSNP households	-1.72	46.6	-0.56	16.0
2010				
PSNP households	-2.03	53.6	-0.44	15.5
Non-PSNP households	-1.79	48.9	-0.42	16.3
2012				
PSNP households	-1.88	50.9	-0.52	16.8
Non-PSNP households	-1.77	47.6	-0.38	15.0

Source: Authors' calculations.

5.3. Impact of PSNP Participation on Children's Nutritional Status

We now consider the impact of PSNP participation on children's nutritional status (Tables 5.6, 5.7, and 5.8).

Recall that the only feasible approach available to us is matching. We use an inverse-probability-weighted regression-adjustment (IPWRA) estimator. We focus on participation in the public works component of the PSNP because nearly all children in our sample are found in that part of the program. We use the following covariates to predict treatment: household livestock holdings 12 months prior to the survey; household landholdings; age, sex, and grade attainment of the household head; the number of males aged 16 to 60 years resident in the household; the number of females aged 16 to 60 years resident in the household; and woreda fixed effects. Our regression adjustment estimates also include characteristics that affect child nutritional status. These include child age and sex; maternal age and schooling; controls for how age, height, and weight were measured; housing quality, and locality characteristics that might affect child nutritional status: whether there was a government health post in the locality where the child resided, whether piped water was available in the locality, and whether the local access road was paved. We experimented extensively with the list of controls; demonstrating that adding or dropping controls does not substantively change the results we obtain. We estimate impacts by survey round, given the substantial differences in program implementation. We estimate these models for all children and for children aged between 6 and 24 months, as the latter are seen to be at particular risk of undernutrition (Hoddinott et al. 2013).

What is striking about Tables 5.6, 5.7, and 5.8 is the near complete absence of statistically significant results. There is no evidence that PSNP participation has any effect on chronic undernutrition as measured by height-for-age z-scores or stunting. No effects are found in any survey round, neither in the full sample nor when we disaggregate by age or sex. These non-results are robust to changes in model specification. We estimated impact models using instrumental variables regressions (instrumenting the real level of payments) and ran woreda fixed effects regressions in which we saturated the model with controls. These did not provide any evidence of impact either. Further, there is no evidence that the PSNP improves acute undernutrition as measured by weight-for-height or wasting. There are some isolated examples of PSNP participation leading to poorer acute nutrition outcomes, but these are few (three out of 24 impact estimates) and only one is statistically significant at the 5 percent level.

Table 5.6 Inverse-probability-weighted regression-adjustment (IPWRA) estimates of the impact of PSNP public works participation on child nutritional status, 2012

	Child age			
	6 to 59 months			6 to 24 months
	All children	Girls	Boys	All children
Height-for-age z-score	0.089	-0.255	0.121	0.260
	(0.110)	(0.155)	(0.154)	(0.201)
Stunting	-0.024	0.033	-0.001	-0.039
	(0.029)	(0.041)	(0.043)	(0.052)
Weight-for-height z-score	-0.191**	-0.137	-0.232	-0.334*
	(0.095)	(0.133)	(0.148)	(0.191)
Wasting	0.033	0.032	0.019	0.050
	(0.020)	(0.027)	(0.032)	(0.040)
Sample size	1,580	797	783	493

Source: Authors' calculations.
Notes: PSNP = Productive Safety Net Programme. Standard errors are in parentheses. *Significant at the 10% level; **significant at the 5% level.

Table 5.7 IPWRA estimates of the impact of PSNP public works participation on child nutritional status, 2010

	Child age			
	6 to 59 months			6 to 24 months
	All children	Girls	Boys	All children
Height-for-age z-score	-0.017	-0.131	0.024	-0.146
	(0.098)	(0.134)	(0.139)	(0.199)
Stunting	0.001	0.027	-0.002	0.080
	(0.028)	(0.037)	(0.038)	(0.049)
Weight-for-height z-score	-0.052	-0.174	-0.006	-0.037
	(0.087)	(0.115)	(0.122)	(0.183)
Wasting	0.010	0.015	-0.002	0.017
	(0.019)	(0.024)	(0.027)	(0.036)
Sample size	1,728	873	855	491

Source: Authors' calculations.
Notes: IPWRA = inverse-probability-weighted regression adjustment; PSNP = Productive Safety Net Programme. Standard errors are in parentheses. *Significant at the 10 percent level; **significant at the 5 percent level.

Table 5.8 IPWRA estimates of the impact of PSNP public works participation on child nutritional status, 2008

	Child age			
	6 to 59 months			6 to 24 months
	All children	Girls	Boys	All children
Height-for-age z-score	0.039	-0.044	0.214	0.148
	(0.126)	(0.166)	(0.177)	(0.231)
Stunting	0.020	0.041	-0.047	-0.049
	(0.038)	(0.047)	(0.050)	(0.055)
Weight-for-height z-score	0.074	0.131	0.038	-0.143
	(0.109)	(0.142)	(0.165)	(0.206)
Wasting	0.016	-0.002	0.026	0.076*
	(0.025)	(0.036)	(0.036)	(0.042)
Sample size	1,338	643	695	385

Source: Authors' calculations.
Notes: IPWRA = inverse-probability-weighted regression adjustment; PSNP = Productive Safety Net Programme. Standard errors are in parentheses. *Significant at the 10 percent level; **significant at the 5 percent level.

5.4. Contextualizing the Non-impact

At the household level, the PSNP has improved food security—see, for example, Gilligan et al. (2009) and Berhane et al. (2015). Why, then, has it had no impact on children's nutritional status? While we cannot provide a definitive explanation, we can suggest some possible reasons.

We begin by examining associations between the PSNP, community-based nutrition (CBN) activities, and children's nutritional status.[9] CBN programs were phased in gradually from 2008 to 2012. We begin by documenting exposure to CBN in these data and its associations with access to health extension workers, exposure to nutrition messages, and one good nutrition practice. We look at the overlap between access to these services, messages, and practices and participation in the PSNP. We then look at the information we have on children's diets. Finally, we look at associations between participation in the PSNP and our measures of children's nutritional status, conditioning on a number of characteristics, including participation in CBN activities.

Table 5.9 shows the number of children in our sample by the year CBN was introduced into the woreda in which they reside and by survey year.[10]

Table 5.9 Number of children in sample, by year of community-based nutrition (CBN) program introduction and survey year

	Survey year			
Year CBN introduced	**2008**	**2010**	**2012**	**All children**
2008	345	294	269	908
2009	–	693	595	2,086
2010	–	510	443	1,482
2011	–	–	247	884
2012	–	–	540	1,927
Number of children in woredas where CBN not yet introduced	2,381	970	–	
Total	2,726	2,467	2,094	7,287

Source: Authors' calculations.

The PSNP surveys were not designed to assess the impact of CBN. However, the 2012 survey round did collect some information on mothers' contact with health extension workers and members of the Women's Development Army[11], as well as questions about exposure to messages regarding foods to feed children under 3 years of age and whether mothers boil drinking water.

Table 5.10 shows that across the full sample in 2012, only 33 percent of mothers had been visited by a health extension worker in the previous month. Just over 15 percent had been visited by someone from the Women's Development Army, and only a quarter had been given information about foods to feed young children. It is rare for households to boil water—less than 12 percent do so. Mothers living in woredas where CBN had been introduced more recently were more likely to have been visited by a health extension worker in the previous month than mothers living in woredas where the program had been introduced earlier.

[9] At the time these surveys were fielded, CBN in Ethiopia had a number of components. These included monthly community sessions to monitor and promote the growth of children 2 years of age and younger, the use of ready-to-use therapeutic foods (RUTFs) to treat severe acute undernutrition, and the counseling of mothers on feeding and other childcare practices.

We include CBN following a specific joint request from the government of Ethiopia and the donors to the PSNP.

[10] Note that this information is not available for all woredas in our sample.

[11] Refers to women community members that are trained by Health Extension Workers in Ethiopia to provide basic health services in the community including immunization and CBN.

Table 5.10 Exposure to nutrition services and messages, by year CBN introduced, 2012

Year CBN introduced	In the last month: Have you been visited by a health extension worker? (%)	Have you been visited by someone from the Women's Development Army? (%)	Have you been given information about foods to feed young children? (%)	Does your household boil drinking water before use? (%)
2008	32.7	11.5	27.6	12.3
2009	35.5	21.7	33.7	13.2
2010	25.1	8.3	23.5	9.5
2011	24.2	13.1	10.2	8.9
2012	40.7	17.7	27.8	11.1
Total	33.0	15.5	26.5	11.2

Source: Authors' calculations.
Note: CBN = community-based nutrition.

Disaggregating these data by PSNP status as of 2012 reveals a similar picture (Table 5.11). Approximately one-third of mothers in 2012 PSNP households had been visited by a health extension worker in the previous month. Just over a quarter, 28 percent, had been given information on foods to feed young children, and only 11 percent boil drinking water. There is no meaningful difference if we disaggregate these data by type of PSNP benefits, public works or direct support.

Table 5.11 Exposure to nutrition services and messages, by PSNP beneficiary status, 2012

PSNP beneficiary status in 2012	In the last month: Have you been visited by a health extension worker? (%)	Have you been visited by someone from the Women's Development Army? (%)	Have you been given information about foods to feed young children? (%)	Does your household boil drinking water before use? (%)
Not a beneficiary	33.4	14.5	28.5	11.2
PSNP beneficiary	33.3	18.2	26.0	11.4
Public works participant	33.5	18.8	26.3	11.6
Direct support recipient	37.4	16.8	28.0	11.2

Source: Authors' calculations.
Note: PSNP = Productive Safety Net Programme.

Next, we examine the data we have on diets of children 6 to 24 months old in the 2012 survey round. Recall that the enumerators had a list of foods (for example, eggs) or food groups (for example, any foods made from beans, peas, lentils, nuts, or seeds). For each item, mothers were asked if their children had consumed that food yesterday during the day or at night. We are especially interested in the extent to which mothers feed foods to their children, other than porridge or *injera* (flatbread) or, given this age group, breast milk. Basic results are shown in Table 5.12.

Table 5.12 Foods consumed by children aged 6 to 24 months during the previous day, all children surveyed, 2012

Region	Any pulses	Any dark leafy vegetables or vitamin A-rich fruits	Other fruits or vegetables	Any milk or other dairy products	Any eggs	Any meat, poultry, or fish	Any fats or oils
Tigray	22.5	14.7	8.5	12.4	20.9	3.9	17.1
Amhara	16.0	16.0	12.3	21.7	7.5	6.6	21.7
Amhara-HVFB	15.5	7.7	4.5	13.5	3.9	5.8	8.4
Oromiya	7.5	14.5	13.7	48.6	9.8	3.5	15.7
SNNPR	4.0	30.5	12.0	37.5	5.5	2.5	15.5
All observations	11.5	17.3	10.7	30.7	9.1	4.1	15.3

Source: Authors' calculations.

What is most striking about the data reported in Table 5.12 is how few children consume animal-source proteins through the consumption of eggs, meat, poultry, or fish; protein or iron through the consumption of pulses; or vitamin A or C through the consumption of dark leafy vegetables or fruit. Across all children, 46 percent consumed none of the foods listed in Table 5.12 and only 11 percent consumed three or more. The most common food consumed is milk or other dairy products, with dairy products consumed most frequently by children living in Oromiya.

Is the consumption of these food items associated with either PSNP participation or the duration of inclusion in the CBN program? Using our IPWRA estimator, we explore this in Table 5.13.

Table 5.13 Impact of PSNP public works participation on child dietary diversity, 2012

	Number of different foods consumed	Yesterday, did the child consume any:				
		Pulses	Fruits or vegetables	Dairy products	Eggs	Fats or oils
Household received PSNP public works payments, 2012, 0/1	-0.002 (0.110)	0.020 (0.029)	0.011 (0.039)	0.008 (0.044)	-0.001 (0.029)	-0.017 (0.036)

Source: Authors' calculations.
Notes: PSNP = Productive Safety Net Programme. Standard errors are in parentheses. *Significant at the 10% level; **significant at the 5% level. Sample size is 543.

The striking feature of Table 5.13 is that participation in the PSNP has no impact on the quality of children's diets. There are no statistically significant relationships between the consumption of any of these foods and the household receiving PSNP benefits in 2012. Adjusting the covariates used as control variables or defining PSNP participation in slightly different ways (for example, based on whether the household received payments for public works or received direct support payments), has no effect on these results.

6. SUMMARY

The PSNP has been successful in improving household food security (Gilligan et al. 2009; Berhane et al. 2015). However, children's nutritional status in the localities where the PSNP operates is poor, with 48 percent of children stunted in 2012. This leads to the question of whether the PSNP could improve child nutrition.

In this paper, we examine the impact of the PSNP on children's nutritional status over the period 2008 to 2012. Undertaking such an exercise requires paying particular attention to the targeting of the PSNP and how payment levels have evolved over time. Using inverse-probability-weighted regression-adjustment estimators, we find no evidence that the PSNP reduces either chronic undernutrition (height-for-age z-scores, stunting) or acute undernutrition (weight-for-height z-scores, wasting). While we cannot provide a definitive reason for this non-result, we note that child diet quality is poor. We find no evidence that the PSNP improves child consumption of pulses, oils, fruits, vegetables, dairy products, or animal-source proteins. Most mothers have not had contact with health extension workers nor have they received information on good feeding practices. Water practices, as captured by the likelihood that mothers boil drinking water, are poor. These findings, along with work by other researchers, have informed revisions to the PSNP. Future research will assess whether these revisions have led to improvements in the diets and anthropometric status of preschool children in Ethiopia.

REFERENCES

Abadie, A., and G.W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74: 235–267.

Berhane, G., J. Hoddinott, N. Kumar, A.S. Taffesse, M. Diressie, Y. Yohannes, R. Sabates-Wheeler, M. Handino, J. Lind, M. Tefera, and F. Simma. 2011. "Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme." Unpublished, International Food Policy Research Institute, Washington, DC.

Berhane, G., S. Gidey, J. Hoddinott, N. Kumar, A.S. Taffesse, Y. Yohannes, J. Lind, M. Handino, K. Roelen, M. Tefera, and B. Nishan. 2013. "The Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme in the Ethiopian Highlands, 2012: Program Performance Report." Unpublished, International Food Policy Research Institute, Washington, DC.

Berhane, G., K. Hirvonen, J. Hoddinott, N. Kumar, A.S. Taffesse, Y. Yohannes, J. Lind, R. Sabates-Wheeler, A. Strickland, M. Tefera, and B. Nishan. 2015. "PSNP-HABP Impact Evaluation, 2014." Unpublished, International Food Policy Research Institute, Washington, DC.

Black, R.E., C.G. Victora, S.P. Walker, Z.A. Bhutta, P. Christian, M. de Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, and R. Martorell. 2013. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *Lancet* 382: 427–451.

Coll-Black, S., D. Gilligan, J. Hoddinott, N. Kumar, A.S. Taffesse, and W. Wiseman. 2012. "Targeting Food Security Interventions in Ethiopia: The Productive Safety Net Programme." In *Food and Agriculture in Ethiopia: Progress and Policy Challenges*, edited by P. Dorosh and S. Rashid. Philadelphia: University of Pennsylvania Press.

CSA (Central Statistical Agency [Ethiopia]) and ICF International. 2012. *Ethiopia Demographic and Health Survey 2011*. Addis Ababa, Ethiopia, and Calverton, MD, US: CSA and ICF International.

de Onis, M., A.W. Onyango, E. Borghi, A. Siyam, C. Nishida, and J. Siekmann. 2007. "Development of a WHO Growth Reference for School-Aged Children and Adolescents." *Bulletin of the World Health Organization* 85 (9): 660–667.

Ethiopia, MoARD (Ministry of Agriculture and Rural Development). 2004. *Productive Safety Net Programme: Programme Implementation Manual*. Addis Ababa, Ethiopia: MoARD.

———. 2009a. *Food Security Programme, 2010–2014*. Addis Ababa, Ethiopia: MoARD.

———. 2009b. *Household Asset Building Programme, 2010–2014*. Addis Ababa, Ethiopia: MoARD.

Gilligan, D., J. Hoddinott, A. Seyoum, S. Dejene, N. Tefera, and Y. Yohannes. 2007. "Ethiopia Food Security Program: Report on 2006 Baseline Survey." Unpublished, International Food Policy Research Institute, Washington, DC.

Gilligan, D., J. Hoddinott, N. Kumar, A.S. Taffesse, S. Dejene, F. Gezahegn, and Y. Yohannes. 2009. "Ethiopia Food Security Program: Report on 2008 Survey." Unpublished, International Food Policy Research Institute, Washington, DC.

Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64(4): 605–654.

———. 1998. "Matching as an Econometric Evaluation Estimator." *Review of Economic Studies* 65(2): 261–294.

Hoddinott, J., J. Maluccio, J. Behrman, R. Martorell, P. Melgar, A.R. Quisumbing, M. Ramirez-Zea, A.D. Stein, and K.M. Yount. 2013. "Adult Consequences of Growth Failure in Early Childhood." *American Journal of Clinical Nutrition* 98: 1170–1178.

Imbens, G.W., and J.M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* (American Economic Association) 47(1): 5–86.

Lemma, F., and J. Matji. 2013. "Delivery Platforms for Sustained Nutrition in Ethiopia." *Lancet* 382: 488–489.

WHO (World Health Organization). 2006. *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height, and Body Mass Index-for-Age: Methods and Development*. Geneva: WHO.

About the Authors

Guush Berhane is a Research Fellow in the Ethiopia Strategy Support Program of the Development Strategy and Governance Division of IFPRI based in Addis Ababa, **John Hoddinott** is from the Division of Nutritional Sciences at Cornell University, and **Neha Kumar** is a Senior Research Fellow in the Poverty, Health, and Nutrition Division of IFPRI based in Washington DC.

Acknowledgment

This work was co-funded by 3ie under grant TW1: 1079 and the Transform Nutrition research consortium, funded by the UK Department for International Development. Preliminary results from this work were presented at four events focusing on Ethiopia: a research workshop, “Safety Nets, PSNP, and Nutrition” in May 2014; a donor briefing in August 2014; the conference “Together for Nutrition” in June 2015; and the 2015 meeting of the Ethiopian Economics Association.

About ESSP

The Ethiopia Strategy Support Program is an initiative to strengthen evidence-based policymaking in Ethiopia in the areas of rural and agricultural development. Facilitated by the International Food Policy Research Institute (IFPRI), ESSP works closely with the government of Ethiopia, the Ethiopian Development Research Institute (EDRI), and other development partners to provide information relevant for the design and implementation of Ethiopia’s agricultural and rural development strategies. For more information, see http://www.ifpri.org/book-757/ourwork/program/ethiopia-strategy-support-program; http://essp.ifpri.info/; or http://www.edri-eth.org/.

About these working papers

The ESSP Working Papers contain preliminary material and research results from IFPRI and/or its partners in Ethiopia. The papers are not subject to a formal peer review. They are circulated in order to stimulate discussion and critical comment.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
2033 K Street, NW | Washington, DC 20006-1002 USA
T: +1.202.862.5600 | F: +1.202.457.4439
Skype: ifprihomeoffice | ifpri@cgiar.org | www.ifpri.org

IFPRI–ESSP ADDIS ABABA
P.O. Box 5689, Addis Ababa, Ethiopia
T: +251.11.617.2000 | F: +251.11.646.2318
mahlet.mekuria@cgiar.org | http://essp.ifpri.info

ETHIOPIAN DEVELOPMENT RESEARCH INSTITUTE
P.O. Box 2479, Addis Ababa, Ethiopia
T: +251.11.550.6066; +251.11.553.8633 | F: +251.11.550.5588
info@edri-eth.org | www.edri-eth.org

The Ethiopia Strategy Support Program (ESSP) is financially supported by the United States Agency for International Development (USAID) and the Department for International Development (DFID) of the government of the United Kingdom and is undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). This publication has been prepared as an output of ESSP and has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and do not necessarily reflect those of IFPRI, the Ethiopian Development Research Institute, USAID, DFID, PIM, or CGIAR.