Direct and Indirect Effects of Malawi's Public Works Program on Food Security*

Kathleen Beegle†, Emanuela Galasso‡, and Jessica Goldberg§

April 12, 2017

Abstract

Labor-intensive public works programs are important social protection tools in low-income settings, intended to supplement the income of poor households and improve public infrastructure. In this evaluation of the Malawi Social Action Fund, an at-scale, government-operated program, across- and within-village randomization is used to estimate effects on food security and use of fertilizer. There is no evidence that the program improves food security and suggestive evidence of negative spillovers to untreated households. These disappointing results hold even under modifications to the design of the program to offer work during the lean rather than harvest season or increase the frequency of payments. These findings stand in contrast to those from large PWPs in India and Ethiopia, and serves as a reminder that public works programs will not always have significant and measurable welfare effects.

JEL Codes: I31, J22, O1. Keywords: public works, food security, Malawi.

*These are the views of the authors and do not reflect those of the World Bank, its Executive Directors, or the countries they represent. This project was funded by the World Bank Research Committee, the Knowledge for Change program, and GLM-LIC. We thank Charles Mandala, John Ng'ambi, and the team at the Local Development Fund in Malawi for their support of the evaluation. We thank Tavneet Suri for her inputs at the design stage of the evaluation. James Mwera and Sidney Brown provided outstanding supervision of field activities. We thank Elizabeth Foster for her brilliant assistance with the data. We are grateful for research assistance from Sai Luo and feedback from Jenny Aker, Arthur Alik, Alejandro de la Fuente, Pascaline Dupas, B. Kelsey Jack, Pamela Jakiela, Sebastian Galiani, Judy Hellerstein, Martin Ravallion, Simone Schaner, Jeffrey Smith, and participants at PAA 2014, NEUDC 2014, IZA GLM-LIC 2015, and seminars at the Inter American Development Bank, the University of Washington, the University of Pennsylvania, Fordham University, and Bocconi University. All errors and omissions are our own.

†World Bank Ghana office. Email kbeegle@worldbank.org.

‡Development Economics Research Group, World Bank, 1818 H St. NW, Washington DC 20433. Email egalasso@worldbank.org.

§3115C Tydings Hall, Department of Economics, University of Maryland, College Park MD 20742. E-mail goldberg@econ.umd.edu.

1 Introduction

Labor-intensive public works programs (PWPs) are common social protection tools in low-income settings (Grosh et al. 2008). These programs require that beneficiaries work in order to receive a cash payment or in-kind transfer (Besley & Coate 1992). They have been widely promoted as tools to protect poor households in the face of large macroeconomic or agroclimatic shocks, due to their relatively rapid rollout (Ravallion 1999). They are recently getting attention in fragile states as tools to quickly restart local economic activities or target the employment of high risk groups (Blattman & Ralston 2015). Well-known examples include the Employment Guarantee Scheme in Maharashtra (Ravallion, Datt & Chaudhuri 1993), the National Rural Employment Guarantee Act (NREGA) in India (Dutta et al. 2014), and the Productive Safety Net Project (PNSP) in Ethiopia (Hoddinott et al. 2012). Such programs are widespread – albeit not on as large a scale – in Sub-Saharan Africa, where 39 of 48 countries have government-supported PWPs (World Bank 2015). They have been increasingly used as a building block of national social protection portfolios.[1]

While many studies of cash-for-work programs focus on the potential crowding out effect of the program on labor market outcomes or the extent of self-targeting for a given wage rate or participation requirement (Alatas et al. 2013, Murgai, Ravallion & vandeWalle Forthcoming), there is surprisingly limited evidence about the first order effects of the programs in increasing consumption levels or allowing beneficiaries to smooth consumption. This paper adds to the literature about the impact of these programs by estimating the effect of Malawi's large-scale PWP, which operates under the Malawi Social Action Fund (MASAF). The stated objectives of the program are to improve food security and to increase the use of fertilizer and other agricultural inputs. Though the PWP increased incomes by offering beneficiaries the opportunity to earn up to US$44 in a country with a per capita gross national income of only US$320, we find no indication that the program achieved its objectives.

Malawi's PWP has been operational since the mid-1990s and provides short-term, labor-intensive employment opportunities to poor, able-bodied households. The implementation of the program is decentralized, with funding allocated to each of Malawi's 31 districts based on population and food security estimates carried out by the government in collaboration with the World Food Programme (WFP). The food security objective is addressed through a combination of support for short-term consumption as well as promotion of medium-term food security through investments in fertilizer, which is intended to increase yields in the subsequent season. Since 2004, the program has been designed to complement Malawi's

[1] Of the 19 Sub-Saharan African countries for which the World Bank's ASPIRE database has cost information, spending on PWPs totals more than 0.1 percent of GDP in nine and accounts for more than 10 percent of the country's social protection budget in eight.

large-scale fertilizer input subsidy program by synchronizing the availability of public works employment with the availability of fertilizer coupons, during planting season. Malawi's PWP ranks fourth in population covered among all such programs in low- and middle-income countries (World Bank 2015).

We use a randomized controlled trial to evaluate the program, and to test the hypothesis that changes to the timing of the program could increase its effect on food security, potentially at the cost of investment in fertilizer. A randomized evaluation of this at-scale program is possible because it is oversubscribed: more villages request PWP activities than can be accommodated given the government's budget, and, even in villages that have projects, not all able-bodied poor households are included. The evaluation includes two levels of randomization: across villages and across households in treated villages. Villages that requested PWP projects were assigned to either a pure control condition (no PWP at all) or one of four treatment groups that offered the same wages and total number of days of employment, but differed in the schedule of work and the frequency of payments.

Our results show that Malawi's PWP was not effective in achieving its aim of improving food security during the 2013 lean season. The program did not increase the use of fertilizer or the ownership of durable goods. We do not find evidence that the program affected prices by injecting cash into the economy. There is also no evidence of labor market tightening induced by reduced labor supply or increased reservation wages, which implies a pure income effect. The failure of the PWP to improve nutrition in either the short run (through consumption support) or longer run (because of increased use of fertilizer) is especially troubling because the MASAF PWP is the largest social protection scheme in one of the world's poorest countries.

The two largest PWPs (the NREGA program in India and the PNSP program in Ethiopia) have been shown to improve some measures of household well-being. By comparison, our findings for the Malawi PWP are very disappointing. Nonetheless, it is important to rigorously study program impacts even when the results are minimal or zero to avoid mischaracterizing the potential of PWPs in developing countries on the basis of only positive findings. The NREGA program in India had some success in stabilizing consumption (Ravi & Engler 2015, Zimmermann 2014). NREGA, and to some extent the PSNP, differs from Malawi's PWP in that it functions as a true insurance program that guarantees employment whenever households need it, for up to 100 days, rather than offering employment in a rationed fashion and only in specific time limited windows of 24 days of work in each of two seasons.

Gilligan, Hoddinott & Taffesse (2009) find modest effects of Ethiopia's PSNP on food security, and Hoddinott et al. (2012) find increases in the use of fertilizer and investments in agriculture only when combined with high levels of payments. When the program is

paired with an explicit strategy for improving agricultural productivity, the impacts are larger. A more recent study of this program found significant improvements in food security for households that participated for multiple years (Berhane et al. 2014). Relative to the MASAF PWP, Ethiopia's PSNP has a longer duration and higher-intensity transfers.

Not only did the Malawian program fail to improve food security for treated households, but there is also some evidence of negative indirect effects on untreated households in villages with the PWP, particularly in the Northern region. In contrast, social protection programs in other settings have been found to have positive spillover effects. Imbert & Papp (2015) and Deiniger & Liu (2013) find evidence of a general equilibrium effect of the employment guarantee scheme in India working through an increase in the casual wage rate, with positive spillover effects for incomes of the poorest households. Angelucci & DeGiorgi (2009) document positive spillover effects of the Oportunidades program in Mexico to households ineligible for the program living in the same villages. Their indirect effects operate through risk sharing; ineligible households are able to consume more through an increase in transfers and loans from family and friends in the community. The low levels of risk sharing we detect in our results are inconsistent with the hypothesis of a crowding-out effect of risk sharing networks as a response to the program.

The remainder of the paper documents the details of the program, the experimental design, and the unexpected impacts. Section 2 describes the program and the design of the evaluation. Section 3 describes the data and outcomes of interest. In Section 4, we outline the empirical strategy and the identification of the parameters of interest. We explain our analytic strategy in Section 4. Then, we discuss the results on national and regional food security in Sections 5 and 6, respectively, and for the use of fertilizer in Section 7. Section 8 discusses potential mechanisms for direct and indirect effects of the program. Section 9 concludes.

2 MASAF program and experimental design

The MASAF PWP has been operational since the mid-1990s and aims to provide short-term labor-intensive activities to poor, able-bodied households for the purpose of enhancing their food security, mainly through increased access to farm inputs at the time of the planting period. The program was designed to be interlinked with Malawi's large-scale fertilizer input subsidy program (currently known as FISP) through the implementation of the PWP in the planting months of the main agricultural season when the FISP distribution also occurs. The premise behind this is that the PWP facilitates poor, credit-constrained households to access subsidized fertilizer. This distinguishes Malawi's program from the more traditional PWP

design of implementation during the lean season.

The MASAF program covers all districts of Malawi through a two-stage targeting approach. In the first stage there is pro-poor geographic targeting and in the second there is a combination of community-based targeting and self-selection of beneficiaries. The amount of funds given to a district is proportional to the district's population and to the poverty rates as well as other measures of vulnerability. District officials then target a subset of extension planning areas (EPAs) based on poverty and vulnerability criteria. Traditional Authorities (TAs) in the EPAs then allocate funds to a subset of selected Group Village Headmen (GVH) who each oversee 3-10 villages. The GVH determines how many households will participate in each village based on available funding; the GVH then works with the village committees in each village to select participating households.

In 2012, as a response to a large currency devaluation, the program was doubled in size and scaled up to cover about 500,000 households per year. The duration of project participation increased from 12 days to 48 days, split in two cycles of 24 days each; the cycles were further divided into two consecutive 12-day waves, and payments are generally made within one or two weeks of the end of each wave. Projects were mostly road rehabilitation or construction, with some afforestation and irrigation projects. The wage rate was 300 Malawian kwacha (MK) per day (US$0.92/day) for a total payment of MK 3,600 for a 12-day wave (US$11.01).

Cycle 1 of the PWP is implemented during the planting season (October to December) to align with the timing of the distribution of the Fertilizer Input Subsidy Program (FISP). Cycle 2 of PWP was designed to take place after harvest, in June and July.

2.1 Experimental design

We use a randomized controlled trial to test variants of the PWP that are budget neutral in terms of direct costs. Villages were stratified by geographic region to improve balance and because geographic and cultural differences between the country's three regions mean that policy makers are particularly attuned to regional differences in program impacts. Villages were randomly assigned (by computer) to one of the four treatment groups or a control condition; households within treatment villages were randomly selected to be offered the program.

2.1.1 Village randomization

The villages in our sampling frame were randomly assigned to one of five groups (see Figure 1). The first of these groups is a control group (Group 0) of villages that were not included in the PWP program in the 2012-2013 Season. Groups 1 through 4 participated in the PWP

in the planting season (Cycle 1 of PWP). These four groups vary in terms of the timing of the second cycle of the program and the schedule of payments in both cycles.

Among PWP-participating villages, two variants are tested:

- *Timing of the program:* PWP is designed to take place for two cycles of 24 days, during planting and during post-harvest seasons. In our evaluation design, we maintain the first cycle at the planting season, and vary the timing of Cycle 2 to take place during the lean season (February-March) instead of during/after the harvest season (May onward). Comparing Groups 1-2 and Groups 3-4 measures the consumption smoothing or buffer role of PWP during the lean season.

- *Schedule of payments:* We introduced a variation in the payment approach from a lump-sum payment made after 12 days of work to a split-payment variant. Under the split-payment alternative, participants are paid three days apart, in five equal installments of MK 720 each.[2] The variation of the payment schedule was motivated by extensive qualitative work done in preparation of the design of this project showing that households treat the lump-sum payments of the PWP differently from income generated through short-term casual labor (day-labor activities referred to as "ganyu" in Malawi). Comparing Groups 1 and 3 to Groups 2 and 4 will allow us to compare whether lump-sum payments alter the patterns of consumption and investment during the planting or lean season.[3]

  Payments in the study districts were facilitated by the research team for the purposes of the evaluation. This was intended to ensure that payments were made without delay, on specific schedules. Administrative payment records confirm that there are no differences in time lag between work and payment across the districts.

The payment schedule may have a differential effect depending on the season. While a lump-sum payment may facilitate investment in a lumpy input in December, split payments may help smooth consumption during the lean season. A lump sum in February may be used

[2]The market price of fertilizer in Malawi at the time of this project was approximately MK 5000 for a 50 kg bag. The national fertilizer subsidy program provided roughly half of households in the country with coupons that allow two bags of fertilizer to be purchased for MK 500 each. Because households face high transaction costs when redeeming their fertilizer coupons, including transportation costs, long wait times, and inflexibility in the days on which fertilizer can be purchased at the government shops, it is substantially more efficient to purchase both bags of subsidized fertilizer at once, for MK 1000 plus transportation costs (which are likely to range between MK 200 and MK 500). While MK 3600 (payment for 12 days of work) more than covers the cost of purchasing two bags of subsidized fertilizer, a single incremental payment of MK 720 does not.

[3]Payments in the study districts were facilitated by the research team for the purposes of the evaluation, with physical delivery of the cash in conjunction with the district officials. The split-payment variant slightly increased the cost of implementation. E-payments, which would entail a small marginal cost of delivery, are under consideration for future rounds of PWP.

for staples as well as temptation goods; divided payments can act as a form of commitment savings that will lead to smoother consumption of staples if people otherwise have exhibit a temptation to spend or have high discount rates even over very short periods of time.

Twenty-eight districts are included in the PWP program. For the evaluation, we randomly selected 12 districts,[4] stratifying by the country's three geographic regions to ensure that the study was representative of the country's population and to motivate analysis of heterogeneous effects in the three distinct regions. Within selected districts, the list of PWP-eligible (pre-screened) villages from the District Council and Traditional Authorities was compared with nationally representative survey data from the Integrated Household Survey (IHS3) collected in 2010 and 2011.

The sampling frame for our analysis corresponds to the overlap between the enumeration areas (EAs) sampled for the IHS3 and the list of communities pre-selected for PWP projects in our 12 districts. This resulted in a total of 182 villages (EAs) to be randomly assigned across our five treatment groups (Figure 1). For the villages selected for treatment, we randomly chose one project in the event that the villages are mapped to two, to have unique village-project pairs. The geographical targeting of the program is reflected in the regional breakdown of the sample (see Table 1), with about one-half of the sample drawn from the Southern region, which has a higher incidence of poverty and food insecurity (Machinjili & Kanyanda 2012). Random assignment was stratified by region.

2.1.2 Household randomization

The second level of randomization is at the household level. This level of randomization improves statistical power in the absence of spillovers, and provides a mechanism for testing for the program's indirect effects on non-participating households in the presence of spillovers. Under the decentralized MASAF program, the GVH identifies households that are offered PWP employment within villages selected for the program. The intention of the program is to target poor households with able-bodied adults. As discussed below, we use the 2010/2011 IHS3 survey as a baseline for this study. By chance, then, it is likely that one or two of the 16 randomly surveyed IHS3 households in our villages will be among those chosen by the GVH for the PWP. We term these households as "village chosen beneficiaries."

For this study, we randomly choose 10 households from the 16 survey households in the village to be offered the program. This strategy is analogous to studying a broad expansion of coverage within villages selected for the program. To ensure that the experimentally induced program offer did not affect the village selection process, the list of randomly selected house-

[4]The 12 districts are Blantyre, Chikwawa, Dowa, Karonga, Lilongwe, Mangochi, Mchinji, Mzimba, Nsanje, Ntchisi, Phalombe, and Zomba.

holds was distributed two weeks after lists of the village chosen beneficiaries were submitted to district councils. We define these randomly chosen households as "top up" households who are "treated" with the PWP program. Some small share of the "top up" households will also be village chosen beneficiaries. Additionally, given the coverage rate of the status quo program, one or two of the untreated IHS3 households were likely to be included in the program through the village selection process.

In summary, our study has three groups of households: treated households in PWP villages (top ups who are randomly offered the program), untreated households in PWP villages,[5] and households in non-PWP communities. By focusing on the random offer, we estimate the intent-to-treat (ITT) effects of the program. By comparing the untreated households in PWP villages to households in non-PWP villages, we are able to measure the indirect (spillover) effects of the program.

3 Data

The data for this study come from five rounds of panel household survey data. The basis for the panel was the Integrated Household Survey 3 (IHS3) fielded in 2010/11 by Malawi's National Statistics Office. The IHS3 is a cross-section of 12,288 households in 768 communities (16 households per community) and has extensive household and agricultural modules.

The 16 IHS3 households in our study villages were interviewed in four additional rounds: before the public works projects started during the planting season (November 2012) after the first cycle, pre-harvest (February 2013), after the lean season cycle, post-harvest (April-May 2013) and finally after the completion of the 2012/13 season (November 2013; see Figure 2). In addition to the household survey data, in terms of monitoring the intervention, we have administrative records which include the dates and amounts of payments and the identities of recipients. These records are used to confirm that beneficiaries received payments in accordance with the days they worked.

Our first survey (before PWP began in all but three villages) is, in effect, a second baseline to complement the IHS3. However, it could be tainted by anticipation effects if households in PWP villages modified their behavior before the program began in expectation of the employment opportunities or other changes it would induce. Twenty-three communities (approximately 13 percent of the communities in which the experiment was implemented) were incorrectly classified as included in the IHS3, and are therefore have no IHS3 data. We will refer to them as the "non-IHS3" sample, no true baseline data are available for this group.[6]

[5] Including a few households selected by their village process and not through randomization.

[6] This reflects the complexities of partnering with the Malawian government to both implement the in-

3.1 Food security measures

We examine food security outcomes using eight indicators and a composite measure. Our measures include (log) per capita food expenditure[7] and (log) per capita food consumption in the last week (including home consumption). Total household calories is computed based on the caloric value of the food consumed. A food consumption score is computed following WFP guidelines and aims to capture both dietary diversity and food frequency; it is the weighted sum of the number of days the household ate foods from eight food groups in the last week.[8] We include a measure of the number of food groups consumed in the last week for seven main groups.[9] We have an indicator for whether the household reported reducing meals in the last seven days.[10] A food security score is constructed according to WFP guidelines and takes on a value of -1, -2, -3, or -4 (higher value indicates greater security).[11] We report a resilience index that is the negative of the World Food Program coping strategy index. Our index is calculated as the negative of the weighted sum of the number of days in the past seven days that households had to reduce the quantity and quality of food consumed; higher values indicate greater food security.[12] Finally, since many of these food security measures are overlapping, we construct a principal components analysis index that includes all eight measures (including the two omitted from the main tables due to space constraints) as a composite food security measure.

tervention and collect nationally representative data. Households in these 23 communities were listed and a sample of 16 households was randomly drawn during our November 2012 survey. For the IHS3 Households that could not be re-interviewed, the team drew a replacement household from the original listing. About 9 percent of households are replacements for the original IHS3 household.

[7]Omitted from the main results tables due to space constraints; available upon request.

[8]The score is calculated based on the sum of weighted number of days in the last week the household ate food from eight food groups: (2 * number of days of cereals, grains, maize grain/flour, millet, sorghum, flour, bread and pasta, roots, tubers, and plantains) + (3 * number of days of nuts and pulses) + (number of days of vegetables) + (4 * number of days of meat, fish, other meat, and eggs) + (number of days of fruits) + (4 * number of days of milk products) + (0.5 * number of days of fats and oils) + (0.5 * number of days of sugar, sugar products, and honey). Spices and condiments are excluded. It has a maximum value of 126.

[9]The seven are described in the previous footnote, with exception of the last group (sugars).

[10]Omitted from the main results tables due to space constraints; available upon request.

[11]The food security score is -1 if in the past seven days, the household reports not worrying about having enough food and reports zero days that they: (a) rely on less preferred and/or less expensive foods, (b) limit portion size at meal-times, (c) reduce number of meals eaten in a day, (d) restrict consumption by adults so that small children may eat, or (e) borrow food, or rely on help from a friend or relative. The food security score is -2 if the household reports that it worried about having enough food and reports zero days for actions a-e. The food security score is -3 if the household reports that it relied on less preferred and/or less expensive foods and b-e are zero. The food security score is -4 if the household reports any days for b-e.

[12]Referring to the five actions described in the previous footnote, the resilience index is the negative of the sum of (a) + (b) + (c) + [3 * (d)] + [2 *(e)]. It has a maximum absolute value of 56.

4 Analysis

We analyze the ITT effect of Malawi's PWP using household-level data from two rounds of post-intervention surveys. Recall that our design includes two levels of randomization: village-level randomization which varied PWP availability and payment structure, and household-level randomization which varied the eligibility to participate within PWP villages. Our main results pool across the four variants of the intervention to estimate the effect of any public works opportunities in one's village and the additional effect of being a treated household within a PWP village. Therefore, we capture the direct effect of PWP availability on treated households, and the indirect effect of the program on untreated households in PWP villages. The indirect effect is important in the context of rationing.[13]

Using data from the lean season (survey round 2), we estimate the equation

$$y_{iv} = \alpha + \beta_1 \text{PWP}_v + \beta_2 \text{PWP}_v * \text{Topup}_i + \Gamma_d + \Theta_t + \epsilon_{iv} \quad (1)$$

where the indicator PWP_v is a village-level indicator for the availability of any PWP program, and Topup_i is a household-level indicator that equals one if the household was randomly selected to be offered ("treated") the program and zero otherwise. The coefficient β_1 captures the indirect effect of the program on untreated households in PWP villages. The coefficient on the interaction term β_2 captures the marginal effect of being offered the opportunity to participate in the program. The sum of the two coefficients β_1 and β_2 captures the total effect of PWP on treated households compared with households with no PWP in their villages. All specifications include district and week-of-interview fixed effects; standard errors are clustered at the village level. As of survey round 2 (February 2013), all villages with PWP programs had completed Cycle 1 (during the planting season).

In half of the PWP villages, Cycle 2 took place in the lean season, after survey round 2 and before survey round 3. Therefore, analyses of the pre-harvest data (survey round 3) estimate the effect of the lean season PWP compared with the standard harvest season program for outcomes just before the harvest takes place.[14] The regression specification becomes

$$y_{iv} = \alpha + \delta_1 \text{Lean}_v + \delta_2 \text{Lean}_v * \text{Topup}_i + \delta_3 \text{Harvest}_v + \delta_4 \text{Harvest}_v * \text{Topup}_i + \Gamma_d + \Theta_t + \epsilon_{iv} \quad (2)$$

The sum $\delta_1 + \delta_2$ is the direct effect of having been offered one 24-day cycle at planting time and a second 24-day cycle during the lean season – on outcomes just before the harvest. The

[13] We capture the indirect effect at the community coverage rate, slightly more than the national coverage rate of 15% since we add 10 top-up households.

[14] Anticipation effects due to the (announced) timing of Cycle 2 could affect survey round 2 outcomes. In estimates available upon request, we show that the effects of the lean and harvest season variants were of similar magnitude in survey round 2.

sum $\delta_3 + \delta_4$ captures the direct effect of one 24-day cycle at planting season. The difference between the two sums is the marginal effect on household outcomes just before harvest of an additional 24 days of potential PWP work during the lean season.

The parameters we estimate are ITT, with identification derived from randomized village and household treatment status, rather than the endogenous participation status of households. Intent-to-treat parameters are policy relevant in that the government can control the coverage rate and the offer of PWP activities, but not households' decisions to take it up.

4.1 Balance

To explore the balance between treatment and control villages in terms of pre-treatment covariates and outcomes, we use the IHS3 data from 2010/11. Although our first round of follow-up survey pre-dates the PWP implementation in all but three villages, the survey was conducted after the intervention was announced in treatment villages. Anticipation of the program may have affected household behaviors.

Using the IHS3 data to examine the balance between the two groups of villages therefore means excluding a subset of villages. So, we first examine the 23 villages that were not included in the IHS3 national household survey in 2010/11 but were included in our sample due to administrative errors. For this subset of villages, we do not have data prior to the annoucement of the intervention. They were randomly assigned to our treatment/groups (six, six, three, three, and five villages to Groups 0-4, respectively). The differences between the IHS3 and non-IHS3 villages reflect a composition effect and have bearing on the external validity of the results, but are orthogonal to treatment assignment. Using our first round of follow-up data, we find that households in the non-IHS3 sample are better off than the IHS3 sample, with better educated household heads, smaller household sizes, and fewer children below the age of 14 (see Table 2).

Unfortunately – and surprisingly, given that randomization was conducted by computer – there is imbalance in pre-program food security at both the village and household levels in the 159 villages for which IHS3 data are available. This is apparent in the estimates of equation (1) for most of the food security measures from the IHS3, reported in Table 3.[15] Untreated households in PWP villages had worse food security than households in control villages according to four of six food security measures (columns 1-4, but not 5 or 6) and the PCA index, but not the number of days of ganyu in the previous week. Treated households fared better than their untreated neighbors by all measures, and statistically better in terms

[15]IHS3 surveys were conducted from March 2010 until March 2011. Balancing tests control for month and year of survey. To control for pre-treatment levels of outcome variables in subsequent regressions given the strong seasonality in these measures, we include the residual of each measure regressed on month and year of survey indicators.

of the number of food groups consumed and the PCA index. However, the treated households (like their untreated neighbors) have worse food security than households in control villages according to four of six food security measures and the PCA index. Three of the differences are statistically significant at the 95 percent confidence level; the point estimates and standard errors for the difference between treated households and households in control villages are displayed below the point estimates in Table 3.

The nature of the imbalance is markedly different in the North, where treated and untreated households have better baseline food security than households in control villages, than in the Central or Southern regions. We report balancing tests for each of the three regions in Appendix Table A1.

The IHS sample is well balanced for a range of non food-security outcomes. Appendix Table A2 reports these tests. Out of 20 coefficients, only one is statistically different from zero. This makes the pre-program differences in food security all the more difficult to understand, but suggests that treated households are not systematically better off than untreated households along multiple dimensions of well-being.

Because of the imbalance, we estimate three different specifications: estimates without baseline (IHS3) covariates from the full sample of 182 villages, estimates without covariates from the subsample of 159 IHS3 villages, and estimates that control for the baseline variable of the outcome variable from the IHS3 villages. This strategy makes clear the extent to which changes in point estimates are due to limiting the sample versus controlling for pre-treatment outcomes. The main tables present results from ANCOVA specifications using the 159 IHS3 villages and including pre-treatment levels of the outcome variables, and other specifications are available in the Appendix.

5 National results

5.1 Lean season

We begin by presenting results from equation (1) which estimates the direct effect of the program on treated households and the indirect effect on untreated households.[16] We report ANCOVA specifications controlling for pre-treatment outcomes.[17] The sum of the coefficients

[16] Our ITT estimates measure the effect of being randomly chosen for inclusion in the PWP program, not of participating in the program. A small number of households were not randomly chosen to be included in the program, but were eligible through village selection procedures. In our specification, these households pooled with the untreated households. One might expect village-chosen beneficiary households to have either better outcomes than untreated households, because of the effect of the program, or worse outcomes, if the poorest households were indeed chosen and the program was not sufficient to offset their relative disadvantage.

[17] Estimates for specifications without the baseline value of the outcome variable are available in the Appendix; Table A3 includes the full sample and Table A4 is restricted to the baseline subsample, which matches

$\beta_1 + \beta_2$ (PWP and $PWP * Top - up$), reported below the coefficient estimates, compares households in treatment villages that were offered PWP to households in control villages.

Panel A of Table 4 displays the effect of the program on food security and labor supply, measured during the lean season (in survey round 2). Higher values indicate better food security.

As expected, treatment increases participation in the PWP. In the month preceding the survey, treated households worked for MASAF an average of 6.2 days more than households in control villages. Untreated households in villages with PWP programs worked 1.5 days more than households in control villages, reflecting the inclusion of some of these households in the PWP through the village selection mechanism despite not being randomly selected. This increase did not crowd out labor supplied to the private market. Treated households did 4.9 days of casual wage labor in the month before the survey, not significantly different than the 4.4 days worked by households in control villages.

Despite participation by treated households and the income from that work, we do not observe improvements in food security. Relative to households in control villages, treated households fare worse by five of the six food security measures, though none of the differences are statistically significant. The effect of the program on the PCA index of food security is close to zero (-0.079, in column seven). The 95 percent confidence interval excludes positive impacts of greater than 0.08 standard deviations relative to the outcome in the control group. Overall, a program designed to improve food security did not: households offered the opportunity to participate in public works in November/December 2012 and January 2013 did not have better food security during the lean season than households in villages without a public works program.

The indirect effect of the program on untreated households in villages with PWP is measured by the village-level indicator for public works, PWP (labeled "Control households in treatment villages" in the tables). This coefficient captures the difference between mean outcomes for households in control villages and households that were not randomly chosen to be offered the PWP program in villages that had the program. Village-chosen beneficiaries who were selected by their communities rather than the randomization procedure are included among these six untreated households in our estimates. We interpret this as a spillover effect of the PWP.

This spillover effect is negative for five of six food security outcomes, including a statistically significant decline in the food security index (0.220). The PCA index also falls relative to the control group, though the reduction of 0.157 points or 0.075 SD (column 6) is not

the sample used in the main analysis. An alternative to ANCOVA specifications is to include the PCA index of food security from the IHS in all of the food security regressions. Results are similar if we control for baseline variation in this way; these results are available upon request.

statistically different from zero. Not only does PWP fail to improve the food security of households randomly offered the program, but there are also some indications that it may reduce food security among their neighbors. [18]

5.2 Pre-harvest period

As designed by the government, the second PWP cycle takes place during the harvest period, beginning in May. This timing is suboptimal if the marginal utility of consumption is higher and the opportunity cost of working lower during the lean season. To determine whether changing the seasonality of the program could improve its effectiveness, our evaluation randomly assigned half of the treated villages to move the second cycle, so that the public works start sooner, in March 2013. Survey round 3 took place after the March-April Cycle 2 and just before harvest. At that time, villages assigned to the lean season treatment had completed work cycles in November-January and March, but those assigned to the standard harvest season treatment had not yet begun their second work cycle. The survey captures food security and other outcomes at the end of the lean season, before the harvest has begun. Estimating equation (2) allows us to test the marginal effect of rescheduling the second cycle of PWP for the lean season instead of the harvest season.

We report national results in Table 4, Panel B. As in the top panel, our main results are from ANCOVA specifications and specifications without controls are available in the Appendix. In addition to the coefficients from equation (2), we report the total direct effect of the lean and harvest season programs ($\delta_1 + \delta_2$ and $\delta_3 + \delta_4$, respectively) and p-values for the tests that the direct effect of the lean season treatment equals the direct effect of the harvest season treatment and that the corresponding indirect effects are equal.

The treatment appears to be implemented correctly: treated households in villages with the lean season PWP report an additional 5.2 days of work for the PWP relative to households in control villages. Treated households in villages with harvest season programs, in contrast, do not have more PWP work in the lean season than households in the control villages. There are no significant effects on the amount of labor supplied to casual wage labor activities, and, reflecting the slack season, the mean number of days worked is lower than reported in the previous round of data collection.

The direct effect of the lean season program on the food security of treated households is small and generally not significantly different from zero. The absolute value of the effect

[18] Jensen & Miller (2011) offer an example of a study that attempts to improve nutrition by reducing the cost of certain foods, but finds zero or negative effects. In that paper, price subsidies for staple goods led some households to substitute toward more expensive calories from meat or fats; so total caloric intake fell, and nutrient intake did not rise. We do not observe an increase in dietary diversity, but note that finding a nutritional elasticity close to zero is not without precedent.

on treated households is small for all outcomes, and there is no consistent pattern to the direction of the effect. Similarly, spillover effects on untreated households in lean season PWP villages are close to zero.

Households in villages assigned to the harvest season treatment (scheduled to work soon after survey 3 was conducted) do not appear to fare as well, though effects are imprecisely estimated. For these treated households, there are declines in five of six food security measures. The PCA index falls by 0.122 or 0.06 standard deviations, not statistically different from zero. The spillover effect of the harvest season program is negative for four of the six reported outcomes and the PCA index, though none of the changes are significantly different from zero.

We cannot reject that the effect of PWP on households treated with the lean season program equals that of households still to be treated with the harvest season PWP work opportunities. This is despite the additional 24 days of work offered to the former group of households. Food security for treated households in villages with lean season PWP (offered 48 days of PWP) is slightly better than households in control villages; for treated households in harvest season PWP villages (offered 24 days as of survey round 3), food security is slightly worse than among households in the control villages. The difference between the direct effect of the lean season and harvest season programs on the PCA index is only 0.04 standard deviations, and the p-value for the test that the direct effects of the two variants on the PCA index are equal is 0.22. Differences in the indirect effects of the two program variants are even less precisely estimated and show no consistent pattern. In short, there is little evidence that changing the timing of Cycle 2 (the second and final 24 days of PWP) is effective in improving food security.

6 Heterogeneous effects by geographic region

Malawi has three distinct geographic regions, which differ in climate, demography, and economic conditions. The Northern region has higher elevation than the rest of the country and lower population density (National Statistical Office 2015). While maize is the main staple crop in all three regions, cash crop production differs. The North grows coffee for export. Tobacco is cultivated in Central and Southern regions, and tea, cotton, and sugar are also grown in the South. Coffee farming utilizes farmers' cooperatives; both estates and smallholders grow tobacco and tea; and cotton and sugar are typically estate-grown crops. Perhaps reflecting the presence of agricultural estates in the South, adults in the South are more likely to be paid employees (45.1 percent) than they are in the Central region (35.9 percent) or the North (23.9 percent) (National Statistical Office 2015).

The South is the poorest region of the country, and has the highest rate of children under age five who are who are too thin for their age (or wasted, a condition that reflects recent nutritional deficit) – 2.72 percent, compared to 1.75 and 1.78 percent in the Central and Northern[19] regions, respectively (National Statistical Office 2015). Some 39.6 percent of adults in the Northern region have completed primary school, compared to 28.5 percent in the Central region and 26.8 percent in the South (National Statistical Office 2015).

Politically, Malawi is organized by districts, which elect Members of Parliament and have appointed District Commissioners. Many government programs, including the PWP we study, are administered by District Councils and there is no explicit coordination between different districts in the same geographic region. Nonetheless, region is extremely salient to politicians and policy makers in Malawi; the National Statistics Office, among other entities, reports many outcomes at the regional level.

While these differences between the three regions do not provide clear predictions that PWP have more positive impacts in one region than another, they do motivate our analysis of heterogeneous effects. The sampling strategy was designed for both national and regional representativeness, and randomization was stratified at the regional level. A consequence of this strategy is that the sample size in the North, the least densely populated region, is small – 339 households in the full sample, and 310 when restricting analysis to households included in the IHS.

In Table 5, we report estimates of a fully interacted version of equation (1) in order to understand whether the program was more effective in some regions than others, and to consider whether regional variation provides any insight about the mechanisms for the effects we find. The formatting of results is as follows: indirect and direct impacts (corresponding to β_1 and β_2 in the pooled specification of equation 1) are reported separately for each region. Then, we report the total effect on treated households separately for each region, followed by p-values for tests that the indirect and direct effects are the same across the three regions. The last panel of the table includes means and standard deviations of outcome variables separately for each of the three regions.[20]

6.1 Lean season

Program participation among treated households is higher in the South (total treatment effect of 7.1 days), followed by the North (6.4 days) and then the Central region (4.9 days), though the differences across regions are not statistically significant (p=0.21, reported just above dependent variable means). Treated households do not reduce their labor supply to

[19] Excluding the island district of Likoma

[20] For specifications without the baseline value of the outcome variable, see Appendix Tables A6 – A8.

the private market in any of the regions; if anything, PWP crowds *in* employment in the local casual labor market in the North, with households increasing the number of days of ganyu.

The program has the most negative impact on food security in the North. Per capita caloric intake falls by 406 calories per adult equivalent, a large and statistically significant reduction that is statistically different from the direct effects on caloric intake in the other two regions. The food consumption score falls by 5.4 points, representing a large and statistically significant reduction (about 0.3 standard deviations relative to the control mean). However, two other measures of food security improve for treated households relative to households in control villages, albeit not significantly. The food security score improves by 0.172 points (0.12 standard deviations) and households are slightly more resilient. Overall, the PCA index falls by 0.41 points or 0.25 standard deviations relative to the control group and the food consumption score falls by 5.4, though given the small sample size in the sparsely-populated North, the change is not significant at standard confidence levels.

The indirect effect of the program in the North is quite pronounced. Untreated households in PWP villages have worse food security than households in control villages for five of six measures; reductions the food consumption score and food security are significant at the 95 percent confidence level, and reductions in per capita food consumption and resilience are significant at the 90 percent level. The PCA index falls by 0.71 points or 0.43 standard deviations relative to households in control villages in the region. That decline is significant at the 95 percent level.

In the Central and Southern regions, the effects of the PWP are much smaller. For treated households, there are no significant changes in any outcomes. The magnitude of the total effect on the PCA index is close to zero in both the Central (-0.025) and Southern (0.080) regions. Indirect effects are negative but not statistically significant for five of six outcomes and the PCA index in the Central region. In the South, point estimates of the indirect effects are close to zero for all outcomes except the number of food groups consumed, which records a marginally significant increase. The small indirect effects in the Central and Southern regions are a contrast to the substantial negative spillovers in the North. The p-value for equal indirect effects on the PCA index is 0.06.[21]

The absence of household crowding out of employment in the local labor market across regions and across different times of the agricultural season in Malawi stands in contrast with the traditional public works literature where households are assumed to have to pay an opportunity cost of participation (due to the foregone income) in public works programs. A

[21] Differences in the direct and indirect effects of PWP persist after controlling for observable baseline characteristics of households that may differ across regions. See Appendix Table A5 for estimates from specifications including the household characteristics reported in Table 2.

detailed analysis of the LSMS-ISA surveys in Tanzaia, Malawi, Ethiopia and Uganda testing for the completeness in labor markets (Dillon, Brummund & Mwabu 2015) finds evidence of a significant labor surplus in Malawi, in contrast with tighter labor markets like in Ethiopia.

6.2 Pre-harvest period

As with the effects on lean season food security, estimating equation (2) on the full sample obscures important regional heterogeneity. We report heterogeneous effects by region in Table 6. Participation in the lean season is high, with assignment to the lean season treatment causing 6.3, 4.2, and 5.7 days of MASAF work in the Northern, Central, and Southern regions respectively. While households treated with the lean season program work fewer days in the Central region than in other parts of the country – mirroring the results for planting season participation – the differences across region are not statistically significant (p=0.37). By design, households treated with the harvest season program report almost no work for MASAF in survey round 3. There is no evidence of crowding out of employment in private labor markets. If anything, there are signs of crowding-in, with households increasing their supply of ganyu labor. This happens in the North for treated households that were not offered the program during the pre-harvest season, and for lean season treatment households in the Central region.

In the North, the direct effect of the lean season program is small, with point estimates that are positive for four of six outcomes and the PCA index. However, the program appears to generate negative spillovers in this region. There are reductions in all six measures of food security, with reductions in the number of food groups consumed and food security that are significant at the 90 percent confidence level. The PCA index falls by 0.49 SD relative to households in the control group, also marginally significant.

For those households in Northern villages who are yet to get the second cycle of 24 days of PWP, outcomes are worse. The direct effect is negative or close to zero for most outcomes and for the PCA index. The indirect effect that was estimated during the lean season persists during the pre-harvest season: the estimated effect is negative for all outcomes, including marginally significant reductions in dietary diversity and increases in both measures of food insecurity. The negative effect on the PCA index is equivalent to 0.54 SDs of the control group, and is statistically significant at the 95 percent confidence level. The negative spillovers to untreated households that plagued the North in round 2 thus persist and deepen in round 3, and occur for both program variants.

In the Central region, the lean season program improves food security for both treated and untreated households in PWP villages. The direct effect is positive for five of six outcomes and the PCA index, which increases by 0.10 SDs though none of the differences relative

to households in the control villages are statistically significant at conventional levels. The pattern and magnitude of indirect effects are similar: food security improves for five measures, including an imprecisely estimated 0.09 SDs of the PCA index.

The impacts for households in the region that have yet to get the second 24-day cycle (harvest season villages) are puzzling. The direct effects are, if anything, negative. Food security falls along four of six measures, though none significantly. The magnitude of the effect on the PCA index is negative but close to zero. The indirect effects, however, are positive. Food security improves for five of six outcomes. The PCA index is 0.21 SD higher for untreated households in villages with PWP than for households in control villages, though the difference is not statistically significant. Since even the treated households did not work in the period leading up to survey round 3 (but had 24 days of work opportunities in November-January and will receive another 24 days beginning in May), it is conceivable that the indirect effects are due to anticipatory behavior by untreated households who are not on the cusp of another 24 day work cycle, relative to treated households whose next work opportunity is pending.

Finally, the lack of effects of PWP in the South that we observed in survey round 2 continues in survey round 3. For the lean season program, the pattern of direct effects is not encouraging: food security fell for five of six outcomes; the PCA index declined by 0.195 points or 0.10 SD, though the change is not statistically significant. The indirect effects are even smaller in magnitude, with the point estimate of the indirect effect on the PCA index close to zero. The harvest season program (villages awaiting their next 24-day work cycle) has negative direct effects on food security on five of six outcomes and reduces the PCA index by 0.12 SDs; none of the changes are significant. The indirect effects are negative but not significant for four of six outcomes. The decline in the PCA index is -0.12 SDs.

For each of the three regions, we test that the direct effect of the lean season program (where treated households were offered 48 days of work as of survey round 3) equals the direct effect of the harvest season program (which offered 24 days of work before survey round 3), and that the corresponding indirect effects are equal. We fail to reject the equality of the two program variants for most outcomes, but the overall pattern of results for food security in all three regions is more favorable – or at least less damaging – for the lean season program (which offered 48 days of PWP) than the harvest season program (as yet offered 24 days of PWP). We do find statistical evidence of differences across the three region, rejecting equal direct effects on the PCA index of either the lean or harvest season treatments, respectively, in all three regions (p=0.08 for the lean season treatment and 0.02 for the harvest treatment) or equal spillovers due to the harvest treatment (p=0.02).

6.3 Interpreting regional differences

The overall pattern of results in Table 5 suggests that the PWP was less effective – or potentially even slightly harmful – during the lean season in the North than in the other regions of the country. The direct effect of participation in the program was negative and significant for two of the seven food security outcomes in the North. Yet, we cannot reject the null hypothesis of equal effects on the PCA index for treated households during the lean season in all three regions (p=0.54, Table 5).

However, for five of the seven food security measures, control households in treatment villages in the North were significantly worse off after the program. And we can reject the null hypothesis of such equal indirect effects on the PCA index (p=0.06, Table 5). Unfortunately, as we discuss in Section 8, it is challenging to identify a mechanism for these negative indirect effects, or why the indirect effects would be different across regions. This result is not due to differences in the distribution of baseline characteristics across the three regions; the regional differences persist, though remain imprecisely estimated for many outcomes, in specifications that interact treatment indicators with baseline characteristics.[22] We have no reason to believe that the program was implemented differently in the North than in other regions; and, if anything, it was better-targeted in the North. With only three regions, we cannot formally test whether differences in aggregate conditions in the regions explain the differences in indirect effects.

Further complicating the interpretation of regional differences, the direct effects of the PWP in the pre-harvest period, as reported in Table 6, are not worse in the North than in the other regions. However, the indirect effects of both the lean and harvest treatment programs are large, negative, and statistically significant in the North, and as reported in the previous section, we reject the null that the indirect effects of the harvest treatment are equal in all three regions.

With a larger sample size that allowed us to exploit aggregate variation at a sub-region (Traditional Authority) level, we might be able to formally test whether regional differences in accessibility, population density or features of local labor, agriculture and goods markets, drive some of these findings. Given the sample for this study, however, we are limited to simply observing that there is no region in which PWP has clear positive effects, and particular caution should be exercised when implementing it in the North.

[22] Results available upon request

7 Use of fertilizer

Complementarities with the fertilizer subsidy scheme drive the design of the program, and increased fertilizer use is a major stated goal of the PWP. In Malawi, fertilizer is applied twice to both maize (the staple crop) and tobacco (the main cash crop). Since survey round 2 captures fertilizer applied during planting season, we estimate equation (1) to investigate the effect of the program on the use of fertilizer. We examine the probability that a household uses any chemical fertilizer during the 2012/13 season; the log of expenditure on fertilizer for the first and second applications; and the log of the quantity of fertilizer used in the first and second applications. For the expenditure and quantity variables, we use the inverse hyperbolic sine transformation to be able to take logs of variables where some observations are zeros. These results are reported in the first five columns of Table 7. The national point estimates of both the direct and indirect effects on any of these outcomes are close to zero, and none of the coefficients are statistically significant.

However, the results reported in Table 8 indicate important heterogeneity by region: while fertilizer use is generally unchanged in the Central and Southern regions, it increases for both treated and untreated households in PWP villages compared to non-PWP villages in the North. In fact, control households in PWP villages use more fertilizer than their directly-treated counterparts, though the difference between the two groups is not statistically significant. They are significantly more likely to use fertilizer (22.3 percentage points, column 5), use significantly larger quantities, and spend significantly more on fertilizer than households in control villages. For the treated households, the quantity of fertilizer used in both applications and expenditures for the second application increase significantly relative to households in control villages. Increased use of fertilizer cannot affect food security during the lean season, but might translate into higher yields several months later.

8 Mechanisms

Malawi's PWP increases potential household income by providing paid work, if not offset by reductions in other labor supply. Despite this, households offered PWP do not have better food security or use more agricultural inputs as a result of the program, and food security either does not improve or worsens among untreated households in PWP villages. We discuss and, when possible, test potential mechanisms for the lack of positive direct effects, and for the negative spillovers in the Northern and Central regions.

8.1 Study design

8.1.1 Low power

Lack of statistical power is one possible explanation for a null main effect. It would be a more plausible explanation had we found positive but imprecisely estimated point estimates, however. Instead, we find predominantly negative effects that are not statistically different from zero. This is true not only at the national level, but also within each region.

We consider the confidence interval for the effect on treated households in order to assess the magnitude of impacts, focusing on the lean season. We can rule out meaningful positive effects of the program: nationally, the upper bound (at the 95 percent confidence level) of the improvement in food security for five of the six individual indicators is less than 0.2 standard deviations of the outcome in the control group; for the sixth, the number of food groups consumed, the upper bound of the confidence interval is 0.22 standard deviations. The standard deviation of the PCA index for food security is 2.08, and the confidence interval for the effect of the PWP on treated households is [-0.29, 0.24].

By region, the upper bound of the 95 percent confidence interval for any individual outcome is 0.34 SDs in the North, 0.33 SDs in the Center, and 0.38 SDs in the South. For the PCA index, the upper bound of the effect is 0.06 SDs in the North, 0.22 SDs in the Center, and 0.24 SDs in the South. Thus, even moderate direct effects are outside the confidence intervals in each region. It does not appear that lack of statistical power explains the lack of positive effects of the program.

8.1.2 Low take-up

The second possibility is that the household-level intervention, which chose households for inclusion in PWP at random, resulted in low take-up and therefore small ITT estimates. The ITT estimates are not biased, but they are – by construction – smaller in absolute value than the TOT effects. Since assignment to the treatment group increases PWP participation, the TOT and ITT effects have the same sign. Therefore, discussion of the TOT does not offer an explanation for coefficients with unanticipated direction of impact.

As designed and implemented by the government, the program is targeted at vulnerable households, which might participate at higher rates than randomly selected households. Across the full study, 57 percent of treated (that is, randomly-selected) households in our study participated in PWP.[23] Since there are considerable within-village spillovers, using household treatment status as an instrument violates the stable unit treatment value assumption (SUTVA) and is not a valid specification. Instead, we can instrument for PWP

[23] See Table A9 for participation rates by round.

participation using village randomization, where the first stage equation is

$$\text{Any PWP participation}_{iv} = \delta_0 + \delta_1 \text{PWP}_v + \Gamma_d + \Theta_t + \nu_{iv} \tag{3}$$

In this specification, the treatment effect incorporates direct and within-village spillover effects; the assumption is that there are no across-village spillovers within our sample, and this is justified by the distance between study villages. Participation is less than 100 % by design: $\delta_1 = 0.34$, and the first stage F-statistic is 186.07. The national TOT parameters reported in Table A10 are larger in magnitude than the weighted average of the direct and indirect effects reported in Table 4, but imprecisely estimated and always in the direction of reducing food security.

Participation varies somewhat by region at 58 percent of treated households in the North, 43 percent in the Center, and 65 percent in the South (see Table A9). Regional TOT estimates are reported in Appendix Table **??** and indicate worse food security for PWP participants in all regions. Low take-up does not explain the lack of impact of PWP on food security.

8.2 Design of PWP

8.2.1 Value of transfer

A key design feature that may contribute to the lack of a direct effect is the low total value of PWP income. The wage rate for the program was set by the government at MK 300 ($0.92) per day, with total possible earnings of MK 14,400 ($44.16). The wage rate is low by international standards, but, in a country with gross national income per capita of $320 ($730, adjusted for purchasing power parity), $44 is non-trivial. The payment for a 12-day work period (the amount disbursed by the government in each pay parade) is equal to the value of mean weekly food consumption at baseline (and more than 1.5 times weekly food expenditure).

Local political constraints made it infeasible to vary the wage rate for this study; so, our experimental design does not allow us to speak to the effect of PWP with higher wages. Despite extensive consumption and expenditure data in our surveys, we are not able to detect increases in any category: food, agricultural inputs or business investments, non-food consumption, or durables. This limitation is shared with many studies on microfinance, which similarly fail to detect the effect of increased cash on household consumption and expenditures.

The lack of beneficial effects of the additional 24 days of work during the lean season also undermines the idea that a more generous program would transform the effects. Treated households in villages selected for the lean season treatment were eligible for 24 additional

days of work in March (for a total of 48 days from November to March). Yet, in surveys conducted the following month, their food security is no better than either households in control villages or treated households in the harvest season villages (offered 24 days from November to January).

8.2.2 Timing

A second hypothesis related to the design of the PWP program concerns timing of the program. First, the program covers periods where the opportunity cost of time is potentially high. Perhaps work on PWP activities crowded out the labor supplied to the household farms or to the wage labor market. We do not have data about hours of work in household agriculture, but note that since survey round 2 is conducted *before* the harvest, any reduction in food security due to reduced future harvests cannot explain the results and would instead exacerbate the zero or negative effect on food security measured later in the year.

As discussed above, we report the effect of the PWP on labor supply in daily wage markets(*ganyu*) in Table 4 (columns 8 and 9). Both the direct and indirect effects of the program are small and not statistically significant at the extensive margin of participation (not shown); at the intensive margin, PWP, if anything, crowds in wage labor, though standard errors are large. It appears that households have an excess supply of labor, a finding consistent with Goldberg (2016).

It also seems unlikely that poor timing vis a vis other government programs explains the results. PWP and the fertilizer subsidy are administered separately and are not perfectly synchronized. The planting season begins earlier in the South than in the North, and the government activated PWP activities earlier in the South. In three study districts, fertilizer subsidy coupon distribution took place between the first and second 12-day waves of PWP activities, and, in the remaining nine districts, fertilizer coupon distribution overlapped with PWP work and payment. The three districts without overlap were Blantyre (South), Dowa (Center), and Karonga (North). The North accounts for a smaller fraction of total population and therefore of our sample; so, discordant timing in one study district represents a larger share of beneficiaries in that region than in the Center or South.

8.2.3 Targeting

Regressive or ineffective targeting potentially explains both lack of direct effects and negative indirect effects. PWP is intended for the able-bodied poor and uses a combination of community wealth ranking exercises and low wages to target the program. In practice, the characteristics of participants may differ from the eligibility criteria because of differences in how local officials select beneficiaries and in the opportunity cost of participation. There are

two types of targeting that may modulate impacts. The first is the selection of village-chosen beneficiaries. As noted above, some untreated households in our study were village-chosen beneficiaries. We examine the correlation between baseline per capita food consumption and participation of these households as an indication of whether the village selection procedures targeted poorer households. Our preferred measure of baseline food security comes from the IHS3 because, unlike round 1 of the evaluation data, the IHS3 data were collected before the intervention was announced. We cannot assess targeting on short-term food security because survey round 1 data were collected after the program was announced and may be affected by anticipated PWP earnings. However, if PWP is used locally in response to short-term shocks, the lag between the IHS3 and program implementation may explain the lack of correlation between time-varying characteristics such as food security and participation.

Nationally, Figure 3 illustrates that there is very correlation between food security in 2010/11 and participation through the village selection process. This may suggest that the village selection process either responds to short-term food security or relies on criteria that are orthogonal to long-term consumption.[24]

As shown in Figure 4, the relationship between food security and PWP participation varies by region. In the North and Central regions, participation of untreated households was uniform across the distribution of baseline food consumption.[25] In the South, households with lower baseline food consumption were, if anything, *less* likely to be chosen by GVH and participate in PWP.

The second type of targeting is self-targeting, captured by participation by randomly selected households from different points of the distribution of baseline food consumption. This mimics an unrationed PWP such as NREGA. Among treated households, the correlation between accepting PWP work and IHS3 food consumption is negative in the North (indicating pro-poor targeting and self-selection of the poorest households), but not in the Central or Southern regions. Though self-targeting seems more prevalent in the North, there is no evidence of displacement of casual wage labor (*ganyu*) in any region.

Overall, Malawi's PWP is rationed and not very well targeted toward the food insecure. Mistargeting could explain the lack of improvement in food security if the program employed people who had lower marginal propensity to consume food, but the geographic heterogeneity in targeting does not seem to explain the regional heterogeneity in results. PWP was, if anything, slightly better targeted in the region where it led to the most pronounced negative spillovers.

[24]The autocorrelation between log per capita food consumption in the IHS3 and survey round 1 in control villages is 0.30. Over shorter horizons, between any two adjacent survey rounds, the autocorrelation in control villages is close to 0.5.

[25]Food consumption data from the IHS3, accounting for seasonality by detrending by week-of-interview.

8.2.4 Project type

PWP activities included road building and tree-planting, which conceptually could have required different effort levels or induced differential selection by participants. While project type is unlikely to explain the lack of direct effect, different project types across the three regions could lead to heterogeneous effects. In fact, the mix of projects was similar across the three regions. We have limited anecdotal information about the day-to-day work activities of beneficiaries, with no evidence of systematic differences by region. As discussed above, PWP activities do not displace wage work in any region. We conclude that any differences in work activities are unlikely to explain regional differences in program impacts.

8.3 Equilibrium effects

8.3.1 Prices

Spillovers could operate through goods markets. If increased demand by treated households drove up the price of food or other goods, the program may have reduced the purchasing power of both treated households and their untreated neighbors. A change in price level has the potential to explain both the lack of positive effects for treated households and negative effects on their untreated neighbors, though differences in market conditions across the three regions would be necessary to explain why the negative spillovers were observed in the North and Central regions, but not the South.

We test whether village-level prices were different in treatment and control villages using a price index constructed from households' reported prices for the five most commonly purchased goods. The specification for village-level differences is

$$y_v = \alpha + \beta \text{PWP}_v + \Gamma_d + \Theta_t + \epsilon_v \tag{4}$$

Note that this specification estimates the effect of PWP on prices at the coverage rate in the experiment. As reported in Table 9, the difference in the price index between treatment and control villages is neither economically nor statistically significant. Nationally, prices in treatment villages are nearly identical – within one percent – to those in control villages. Neither is there any evidence of price effects at the regional level. In the North, the magnitude of the coefficient on the PWP indicator is near zero; prices are about seven percent lower for treatment villages than controls in the Center and seven percent higher in the South. None of the differences are statistically significant. This likely reflects the low coverage rate of PWP.

8.3.2 Labor markets

Imbert & Papp (2015) find that the large Indian PWP increases wages in the low-skill market. A similar effect on equilibrium wages in Malawi could explain negative externalities by driving up labor costs for untreated households that hire workers, especially since planting season PWP activities coincide with peak labor demand. There was not substitution away from work in the private wage market in Malawi, but the PWP program could have driven up reservation wages.

We did not collect data about wages paid for hired labor. To test the labor market tightening hypothesis indirectly, we interact the PWP treatment variables with an indicator for whether the household reported hiring any agricultural labor in the IHS3; about 20 percent of households had previously hired labor.

We estimate

$$\begin{aligned} y_{iv} =& \alpha + \psi_1 \text{PWP, no hired labor}_i v + \psi_2 \text{PWP, hired labor}_i v \\ &+ \psi_3 \text{PWP, no hired labor}_i v * \text{Topup}_i + \psi_4 \text{PWP, hired labor}_i v * \text{Topup}_i \\ &+ \psi_5 \text{Any previous hired labor}_i + \Gamma_d + \Theta_t + \epsilon_{iv} \end{aligned} \tag{5}$$

ψ_1 is the indirect effect of PWP on households that did not hire any labor during the 2010/11 growing season, and ψ_2 is the indirect effect on households that did. If labor market tightening were responsible for negative spillover effects, we would expect ψ_2 to be negative (and more negative than ψ_1).

Instead, as reported in Table 10, the indirect effect of PWP is negative for households that do not hire, a group that might benefit from, and should not be harmed by, higher equilibrium wages. For the PCA index (column 9), the point estimate of ψ_1 is -0.16, not statistically different from zero at conventional levels but substantial relative to the mean in the control group (0.15). The indirect effect on households that previously hired labor and would face higher costs due to an increase in wages is positive though smaller in magnitude: $\hat{\psi_2} = 0.08$. This test does not support the labor market tightening channel (via wage increases) for the negative spillovers.

8.4 Changes in transfers from institutions or individuals

Other institutions – governmental or nongovernmental programs – or individuals could have changed their behavior towards untreated households or otherwise compensated control villages with other programs. In this case, the untreated households in PWP villages, which benefited from neither PWP nor these alternative programs, may be worse in comparison. We have data about the receipt of five types of benefits: food assistance or maize distribution,

school-based nutrition programs, scholarships or other help with school fees, cash transfers, and fertilizer coupons. As shown in Table 11, there is no evidence that untreated households differ in their receipt of these benefits from either treated households or households in control villages. (Indeed, the only economically or statistically meaningful difference is in fertilizer coupons, which are more likely to be available to treated households in accordance with the designed linkage between PWP and the national fertilizer subsidy scheme.)

A related hypothesis is that treated households were not the final beneficiaries of the program, but instead shared their income with other households or used it to replace or repay loans. We examine four outcomes in the right hand side of Table 11: whether the household received any transfers, whether it made any transfers to others, whether it received any new loans, and whether it made any loan payments in the previous month. Both treated households and untreated households in PWP villages give fewer transfers than their counterparts in control villages, which works against explaining the main findings. PWP does not have a significant direct effect on other transfer outcomes, and the magnitudes of the direct effects are small as well as imprecisely estimated.

It is also possible that contacts reduced transfers to or increased demands for transfers from households in villages with PWP. Even untreated households in PWP villages could be affected through this channel. For example, family members living in other locations may know that PWP was active in a village, but have inaccurate information about who actually benefited from the program. They may falsely believe that untreated households received income from PWP and therefore do not need support through social networks or may even have funds they could share with others. The results reported in columns (6) through (9) of Table 11 indicate that untreated households are significantly *less* likely than households in control villages to make transfers to others, ruling out the possibility of misguided pressure to share income. Untreated households do not significantly differ from controls (or treated households) in their probabilities of receiving transfers or accessing loans, and the effects are not different across the three regions.

8.5 Behavioral responses by treated and untreated households

Finally, households offered PWP may make choices that offset or limit the effect on food security. If households behave according to the predictions of the permanent income hypothesis (PIH), then the temporary increase in income will be allocated to consumption in all future periods, and the change in consumption in any one period would be diminishingly small, as in Chen, Mu & Ravallion (2009). This could account for the lack of positive effects on the food security of treated households, but cannot explain negative spillovers to untreated households. Perhaps more importantly, the myriad credit market imperfections in Malawi

that lead to well-documented seasonal variation in consumption make it unlikely that the PIH holds in this context.

Another possibility is that treated households responded to the program by voluntarily reducing their food consumption to pool existing resources and wages from the program and make indivisible purchases of durables. We test this hypothesis by estimating the effect of PWP on durable ownership (Table 7, columns 6-8). We consider three measures of durable ownership: the probability of owning at least one of 32 goods covered in the survey, the total number of durable goods owned, and a PCA index of durable ownership. The total effect of PWP on treated households is not significantly different from zero for any of the three outcomes, and the imprecise estimates for the number of goods and PCA index are negative. There is no evidence that households offered PWP use program earnings or other resources to purchase durable goods, and these effects do not differ by region (Table 8).

Recall both treated and spillover control households in the North *do* increase their use of fertilizer. The negative direct and indirect effects on food security are also most pronounced in the North. Households in that region may have devoted PWP income and other resources to the purchase of fertilizer rather than the purchase of food. We highlight this result because it is the only evidence we have of behavior that is consistent with the unexpected negative spillovers on food security and matches the geographic pattern of direct and indirect results.

8.6 Remaining possibilities

While the PWP increased income for treated households, it may simply have been too small or too poorly timed to measurably improve their food security or to increase their use of fertilizer. An explanation for the secondary finding of negative spillovers to untreated households is more elusive. Here, we discuss the remaining possibilities to complete our consideration of our findings.

One possible explanation for negative spillover effects is strategic self-reporting by untreated households that do not experience real reductions in food security, but choose to underreport consumption to signal their potential eligibility and increase their chances of being offered PWP or other benefits in the future. If strategic misreporting were responsible for finding negative spillovers, then perhaps it also biased the estimate of the main effect. We argue that two features of our research design make it unlikely. First, data collection and program implementation were conducted by different institutions. The NSO was responsible for survey work, and the Local Development Fund, with assistance from Innovations for Poverty Action, managed PWP activities and payments. Respondents did not report their food consumption or other outcomes to people who were employed by or connected to the administration of the PWP. Second, the IHS3 households in our study were part of a na-

tionally representative survey that is not connected to the provision of any public or private benefits. These households had been administered the same survey instrument by the same institution (NSO) before the study began and had not received any benefits that depended on their answers. These features distinguish our evaluation from randomized controlled trials that rely on data collected by employees of the same organization that administers the intervention, and from evaluations of subjects who may infer that surveys are collected for the purpose of allocating benefits.

We are not able to assess the reliability of our outcome measures to determine whether perceived or self-reported food security may have fallen, absent a change in objective consumption. While the eight different measures of nutrition we examine are all related to food security, they measure different aspects of the concept. Therefore, we cannot use the correlations between items in the IHS3 and subsequent survey rounds as an indication of item reliability, because the program itself may have affected some measures of food security differently than others.

We cannot rule out that PWP changed the behavior of untreated households in some way that worsened food security. Such a mechanism would have to take a specific form to explain both the main effect and the indirect effects, and be consistent with other behaviors we observe. One possibility is as follows: PWP eligibility causes a change in the behavior of treated households that is not captured by our survey (either because it falls outside the recall period, or because it is too small to be measured precisely). Untreated households respond to the behavior or consequences of the behavior by the treated households. For example, treated households might spend their earnings by increasing consumption immediately after payment, which would not be captured by a the seven-day recall period of the survey that takes place a few weeks after the payment. Untreated households respond by similarly increasing short-run consumption to keep up with their neighbors or because their reference point was affected, but then have to reduce consumption in subsequent periods because their short-term splurge was not financed by PWP earnings. This explanation requires peculiar behavior by the untreated households, however: their marginal utility of keeping up with their neighbors is high enough to offset reduced consumption later on, or their preferences are time inconsistent, or they have erroneous expectations about future income. This sort of conspicuous consumption with negative welfare consequences has been documented among neighbors of Canadian lottery winners (Agarwal, Mikhed & Scholnick 2016).

An alternative explanation requiring less complicated behavior by the untreated households is that, immediately after they are paid, treated households purchase food in bulk quantities for storage and later consumption. These purchases are not captured in the survey recall period and therefore do not translate into increased food security from the program, but they do reduce the availability of commodities at local markets for untreated households.

The equilibrium effect is on the quantity of goods, not prices, perhaps because of frictions in the wholesale market. Food security for untreated households falls because of supply-side factors.

We see some evidence of this behavior in the fertilizer purchases of untreated households in the North. Treated and untreated households increase their expenditure on and use of fertilizer by about the same amount – nearly double – relative to controls. For treated households, the MK 6,000 increase in spending across the basal and top dressing fertilizer applications could have come entirely from PWP earnings. Untreated households, though, did not have extra income to use for the fertilizer. They did not have the opportunity to earn wages from public works activities and did not work more in the casual wage labor market. If the increase in the use of fertilizer had not been accompanied by a reduction in food security, we would suspect that treated households had shared the spoils of the program with their untreated neighbors. We see no evidence of this in our direct measures of cash and in-kind transfers. The pattern of food security and fertilizer use results makes it more likely that instead, untreated households in the North modeled their behavior on that of their treated neighbors. We cannot distinguish between an adjustment of the reference point about appropriate fertilizer use or incorrect expectations that PWP or other opportunities would be available to untreated households in the future.

Bazzi, Sumatro & Suryahadi (2015) find that a large-scale conditional cash transfer in Indonesia does not affect consumption growth among households that receive transfers on schedule, but reduces consumption growth among households whose second transfer is delayed. In our context, untreated households in PWP villages may behave like Indonesian households whose transfers were delayed: their expectations about future income changed, but were unfulfilled (at least within the survey recall period).

8.7 The benefits of hindsight

We collected four rounds of comprehensive household survey data and use data from a nationally representative survey as a baseline, and integrated extensive process monitoring throughout the project. However, there are four types of data we did not collect that, in hindsight, could have provided additional insight into the impacts of the PWP and the mechanisms behind those impacts. First, financial diaries data may have been better-suited to understanding the nuances of expenditures that include small purchases of many different goods (Collins et al. 2009, Zollmann 2014, Zollmann & Sanford 2016). The financial diaries approach focuses on cash flows among a small number of households that are not statistically representative of particular populations, and more flexible quantitative data collection not tied to a linear interview. These data would be useful even if they were collected for a

stratified subsample of the population, and in a multi-round project like this, results from financial diaries collected early in the project could have been used to adjust the questionnaires used in subsequent survey rounds. Because our process monitoring leaves us confident that respondents received cash, data that would help us understand where it went are the top priority.

We think concerns about strategic misreporting are unlikely to explain our findings because of the separation between data collection and program implementation and because it is commonly known in Malawi that villages do not participate in the PWP in successive years. That said, the second type of data that could help alleviate remaining concerns include objective measures of physical wellbeing or other outcomes.

For example, data about soil nitrogen levels could be used as a proxy for fertilizer application. Anthropometric data (weight, mid upper-arm circumference, or even biomarkers such as hemoglobin levels) could have provided a direct test of the effect of the program on physical wellbeing (Frankenberg, Ho & Thomas 2016). These data are not subject to concerns about recall or reporting error, and can capture outcomes that are hard to observe from household survey data when resources are not divided equally or predictably among household members. Anthropometric data can be measured quickly and accurately with proper training, but they are also notoriously slow to change. In assessing the benefits of large scale programs like PWPs, a reasonable strategy to contain costs and maximize power might involve collecting outcomes for individuals who are expected to be most sensitive to short-term changes in food availability, such as acute malnutrition in young children.

Haushofer & Shapiro (2016) complement self-reported wellbeing by measuring the stress hormone cortisol. This could offset concerns about strategic misreporting of wellbeing, though stress is only one mechanism through which PWPs could affect happiness.

Third, measures of subjective wellbeing would have been helpful in interpreting the results. Questions on mental health and psychological distress as well as expectation about the future could help to determine whether benefits accrued through a different channel than consumption. A decrease in subjective wellbeing would, of course, support the conclusion that the program does not make participants better off.

Fourth, conducting the household surveys closer to the payment day, such that the survey recall period included the day of the payment, would detect very short-run changes in expenditures and consumption and rule out the "keeping up with the Joneses" hypothesis discussed in the previous section. While this would rule out one possible explanation for the results we obtain, it is not clear that outcomes only at the moment of payment are of interest when evaluating a program with infrequent payments that is nonetheless designed to smooth consumption. A comprehensive data collection strategy might include weekly consumption and time-use surveys, in order to measure contemporaneous and post-program effects, and

to improve statistical power with highly correlated outcomes (McKenzie 2012).

Finally, leveraging nationally representative survey data increased the statistical power for this project and provides a public good: panel data that we expect to be of use to other researchers. Such a strategy is logistically demanding and required that we cede control of some of the day-to-day aspects of data collection to the National Statistics Office. In our case, this led to delays and missing baseline data for some villages. We consider the tradeoff worthwhile, both because of the rich data available to this and other projects, and because of the separation between the implementation and data collection associated with this evaluation.

9 Conclusion

PWPs can stabilize the income and improve the food security of beneficiaries by providing earnings opportunities, and can achieve targeting through low wages and work requirements that promote self-selection. While Malawi's PWP offers households the opportunity to earn approximately $22 at planting season and an additional $22 later in the year, it does not have a measurable short-term effect on lean season food security. Even improving the structure of the program by rescheduling the second work cycle from the harvest season to the lean season does not generate measurable improvements in the food security of treated households. These findings stand in contrast to those from large PWPs in India and Ethiopia, and serves as a reminder that PWPs will not always have significant and measurable welfare effects.

While the maximum possible income from PWP is substantial in light of the country's per capita gross national income of $320, the magnitude of the transfer is low compared with other social protection tools such, as cash transfers being studied in Malawi,[26] or other PWP programs in Sub-Saharan Africa.[27] Our results do not speak directly to the effect of a more generous program, though comparison with results in other countries suggests potential earnings as a margin for increasing the impact of the program. Perhaps because of the low daily wage in MASAF's PWP, 24 extra days of work during the lean season does not significantly improve food security, but longer duration and more flexible schedules are avenues for future investigation.

Households may have spread consumption across the four- to eight-month (depending on

[26] In their study of the cash transfer project in one district of Malawi in 2007, Miller, Tsoka & Reichert (2011) find large, positive effects on beneficiary households in program villages compared with households in control villages screened as eligible but not given the program. In this program, the size of the benefit is significantly larger, with transfers totaling $168 per household over the course of a year (equivalent to about $250 in 2012 price levels), an amount more than five times what households received from PWP in 2012/13.

[27] Comparable PWP programs in the region, such as the ones in Ethiopia, Sierra Leone, and Ghana, have maximum earnings of between $60 and $190.

treatment group) PWP period or saved for even longer durations. Then, changes in weekly spending (the interval captured in our survey period) may be too small to detect, especially since extra spending may have been spread across many different categories of goods. We are confident that participating households did receive the cash, as the project included monitoring of payments issued by district officials and other extensive supervision of the government-managed process. Despite this degree of certainty about payment delivery, we are not able to "find the cash" in the data, so to speak. In hindsight, the project would have benefited from complementing our comprehensive quantitative household surveys with a more nuanced approach to tracking financial flows, such as using financial diaries.

With the available data, we can rule out significant improvements in the two outcomes specifically targeted by the program, food security and the use of fertilizer, but there may have been small, diffuse increases in these or other outcomes that we do not detect. This interpretation allows for the possibility that the PWP was welfare-improving for households that chose to participate and thus is consistent with their revealed preference to participate, while still ineffective in achieving its main policy objectives.

In Malawi, the PWP is designed with an additional goal: it is timed to coincide with the planting season to promote take-up of the country's fertilizer subsidy scheme. However, our results do not support the hypothesis that the two programs are complementary. While households included in PWP are more likely to receive fertilizer coupons (consistent with the policy of interlinkage with the fertilizer subsidy) and hence pay less for the fertilizer they use, they do not use more fertilizer. In the Central and Southern regions, treated households do not increase fertilizer use relative to households in control villages, and, in the North, treated and untreated households in PWP villages both increase their fertilizer use relative to households in control villages. This may suggest that other constraints, including lack of knowledge of the correct amount of fertilizer to use (Duflo, Kremer & Robinson 2008), are more important than credit constraints in limiting the use of fertilizer.

The program is rationed nationally, with funding available to cover only 15 percent of households. When coverage is extended to additional households selected randomly rather than through the community wealth ranking exercises, take-up is about 50 percent. However, participation is higher among less food-secure households only in the North, suggesting that the program is not achieving pro-poor targeting (as measured by these indicators) through self-selection at the current wage rate. Despite this, the program does not displace labor supplied to household farms or casual day labor, likely because of slack labor markets even during peak agricultural periods.

The indirect effects of the PWP are small or, surprisingly, negative. In Northern and Central Malawi, food security of untreated households in villages with PWP programs is not only lower than food security among their treated neighbors, but also lower than food

security in control villages without PWP activities. This is in contrast to expectations and to the effects of other large-scale transfer programs. For example, Oportunidades, the conditional cash transfer program in Mexico, generated positive effects on the consumption of treated households and positive externalities to non-beneficiary households (Angelucci & DeGiorgi 2009).

An explanation for this unexpected finding has proven elusive. Cash transfers in Mexico generated equilibrium effects on commodity prices (Cunha, DeGiorgi & Jayachandran 2014), but we find no evidence of price increases in Malawian villages with PWP. The large PWP in India raised equilibrium wages, but we reject the labor market tightening explanation for negative spillovers in Malawi. Pressure to share money could have explained the negative effect on untreated households if relatives mistakenly believed that, because PWP was present in the village, even the untreated households had benefited and could contribute to the social network. However, there is no evidence of increased income sharing from PWP villages, and, if anything, untreated households made fewer contributions to their networks than households in control villages.

A possibility we cannot exclude is that untreated households reduce food consumption in reaction to an unobserved change in the behavior of treated households or to erroneous expectations of their own future income. We discuss two examples of this: untreated households overspend to match a short-term and thus, unobserved (by the econometrician) increase in consumption by their treated neighbors, and compensate by reducing consumption during the period we do observe, or a combination of stock-outs and sticky prices reduces food availability. The matching and substantial increases in fertilizer use by both treated and untreated households in the North, the region with the strongest negative spillovers, are consistent with this hypothesis. We lack direct evidence to test related behaviors, but present them as examples of the types of mechanisms that could explain the unexpected finding that Malawi's PWP reduces the food security of untreated households in villages with PWP activities. Identifying the mechanism remains a priority for both understanding household spending patterns and informing policy.

Figures

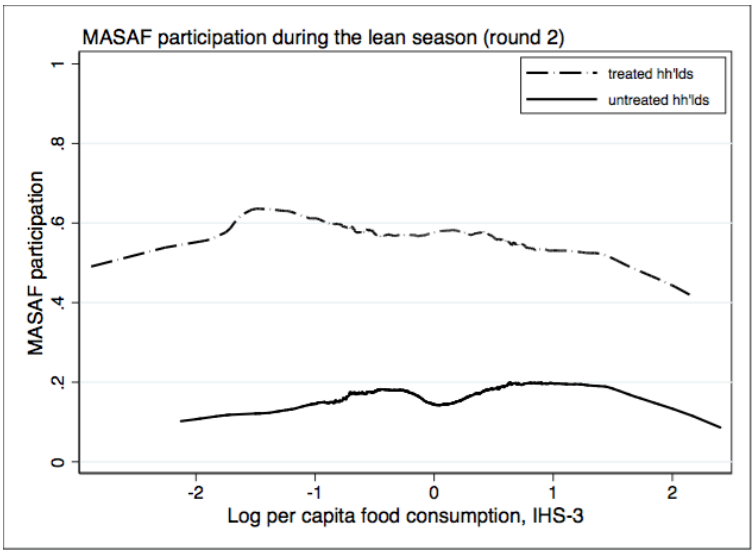
Figure 1: Experimental design

	Treatment at the community level	
	Cycle 1: Planting season PWP Cycle 2: Harvest season PWP	Cycle 1: Planting season PWP Cycle 2: Lean season PWP
No PWP	Group 0 [38 communities]	
PWP with lump sum payment	Group 1 (status quo) [40 communities]	Group 3 [35 communities]
PWP with split payment	Group 2 [34 communities]	Group 4 [35 communities]

Figure 2: Timeline

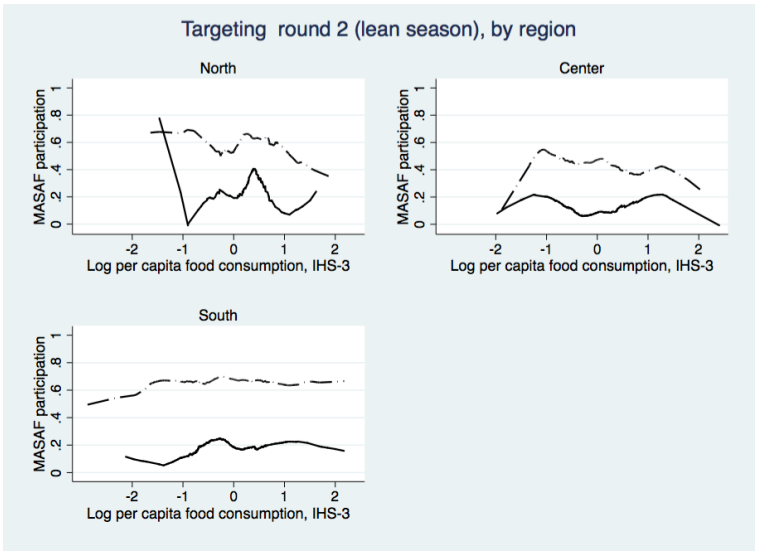
Year-round	Pre-planting	Planting	Lean	Lean	Early harvest	Harvest	Pre-planting
Mar. 2010-Mar. 2011	Oct.-Nov. 2012	Wave 1: Nov.-Dec. Wave 2: Jan. 2013	Feb. 2013	Wave 1: Feb.-Mar. Wave 2: Mar.-Apr.	Apr.-May 2013	Wave 1: May-Jun. Wave 2: Jul.-Aug.	Oct. 2013
IHS3	Survey round 1 (post announcement of treatment)	PWP Group 1 PWP Group 2 PWP Group 3 PWP Group 4	Survey round 2	PWP Group 3 PWP Group 4	Survey round 3	PWP Group 1 PWP Group 2	Survey round 4

Figure 3: National targeting based on IHS3 food security



MASAF participation is an indicator for whether any member of the household has participated in PWP in the month preceding survey round 2 (dashed line is for treated households, and solid line for untreated households. Log per capita food consumption from IHS3.

Figure 4: Regional targeting based on IHS3 food security



MASAF participation is an indicator for whether any member of the household has participated in PWP in the month preceding survey round 2 (dashed line is for treated households, and solid line for untreated households. Log per capita food consumption from IHS3.

Tables

Table 1: Regional breakdown of treatment assignment

	Control community	PWP community	Total
North	4	18	22
	10.5%	12.5%	12.1%
Central	14	50	64
	36.8%	34.7%	35.2%
South	20	76	96
	52.6%	52.8%	52.7%
Total	38	144	182
	100%	100%	100%

Table 2: Descriptive statistics of households (survey round 1) by IHS3 status

	Non IHS3 communities	IHS3 communities	p-value: non IHS3 = IHS3
Female headed household	0.304	0.260	0.800
	(0.461)	(0.439)	
Highest education of HH head	6.337	6.191	0.066
	(3.831)	(3.334)	
Head attended secondary school	0.230	0.186	0.003
	(0.422)	(0.389)	
Household size	4.380	4.969	0.000
	(1.998)	(2.291)	
Number of children under 14	2.149	2.388	0.005
	(1.568)	(1.718)	
Number of EAs	23	159	

Standard deviations in parentheses.

Table 3: Balance tests (IHS3)

Dependent variable:	(1) Ganyu days	(2) Ln (p.c. food cons.) last week	(3) Food cons. score	(4) Per adult equivalent calories	(5) # food groups consumed	(6) Food security score	(7) Resilience index	(8) PCA index
Control household in treatment village	6.386	-0.128**	-3.864**	-120.567*	-0.380***	0.041	0.712	-0.369**
	(5.237)	(0.051)	(1.279)	(68.218)	(0.098)	(0.097)	(0.586)	(0.151)
Treatment household	-2.663	0.047	1.345*	39.926	0.170**	0.118*	0.112	0.188**
	(4.084)	(0.034)	(0.780)	(54.561)	(0.062)	(0.062)	(0.288)	(0.085)
Total effect on treatment households (relative to control villages)	3.723	-0.082*	-2.519**	-80.641	-0.210**	0.158*	0.823	-0.181
	(5.032)	(0.046)	(1.199)	(62.543)	(0.091)	(0.094)	(0.591)	(0.146)
Mean of dep. var. in control villages	37.73	6.16	45.33	2546.25	5.08	-2.34	-4.57	0.06
S.D. of dep. var. in control villages	72.42	0.70	15.51	1059.84	1.24	1.43	8.15	1.88
R-squared	0.16	0.15	0.12	0.14	0.13	0.23	0.21	0.18
Observations	2274	2274	2270	2274	2270	2274	2274	2256

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
Estimates include district and week-of-interview fixed effects.
Ganyu days (column 1) is reported for the last month. We examine food security outcomes using six indicators and a composite measure. Log per capita food consumption (column 2) is for the last week and includes home consumption. A food consumption score (column 3) is computed following WFP guidelines and aims to capture both dietary diversity and food frequency; it is the weighted sum of the number of days the household ate foods from eight food groups in the last week. The score is calculated based on the sum of weighted number of days in the last week the household ate food from eight food groups: (2 * number of days of cereals, grains, maize grain/flour, millet, sorghum, flour, bread and pasta, roots, tubers, and plantains) + (3 * number of days of nuts and pulses) + (number of days of vegetables) + (4 * number of days of meat, fish, other meat, and eggs) + (number of days of fruits) + (4 * number of days of milk products) + (0.5 * number of days of fats and oils) + (0.5 * number of days of sugar, sugar products, and honey). Spices and condiments are excluded. It has a maximum value of 126.
Per adult equivalent calories (column 4) is computed based on the caloric value of the food consumed, and is winsorized at the 10th and 90th percentiles. We include a measure of the number of food groups consumed (column 5) in the last week for seven main groups. The seven are described in the construction of food consumption score, with exception of the last group (sugars). A food security score (column 6) is constructed according to WFP guidelines and takes on a value of -1, -2, -3, or -4 (a lower absolute value indicates greater security). The food security score is -1 if in the past seven days, the household reports not worrying about having enough food and reports zero days that they: (a) rely on less preferred and/or less expensive foods, (b) limit portion size at meal-times, (c) reduce number of meals eaten in a day, (d) restrict consumption by adults so that small children may eat, or (e) borrow food, or rely on help from a friend or relative. The food security score is -2 if the household reports that it worried about having enough food and reports zero days for actions a-e. The food security score is -3 if the household reports that it relied on less preferred and/or less expensive foods and b-e are zero. The food security score is -4 if the household reports any days for b-e. Larger values indicate improved food security. For simplicity and to maintain consistent interpretation of the sign of the outcome variables, we report a resilience index (column 7) that is the negative of the World Food Program coping strategy index. Our index is calculated as the negative of the weighted sum of the number of days in the past seven days that households had to reduce the quantity and quality of food consumed. Referring to the five actions described in the construction of food security score, the coping strategy index is the negative of the sum of (a) + (b) + (c) + [3 * (d)] + [2 *(e)]. It has a maximum absolute value of 56. Finally, since many of these food security measures are overlapping, we construct a principal components analysis index (column 8) that includes eight measures (including the two omitted from the main tables due to space constraints) as a composite food security measure. The two omitted measures are log per capita food expenditure, and an indicator for whether the household reported reducing meals in the last seven days. Higher values indicate better food security in column 2-8.

Table 4: ITT Effects on Food Security and Labor Supply (Rounds 2 and 3)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: Round 2									
Control household in treatment village	1.589***	-0.084	-0.026	-1.071	-42.518	0.067	-0.212**	-0.089	-0.108
	(0.333)	(0.779)	(0.048)	(1.098)	(71.233)	(0.113)	(0.082)	(0.660)	(0.145)
Treatment household	4.667***	0.488	0.014	0.363	-19.910	0.046	0.151**	-0.135	0.079
	(0.338)	(0.465)	(0.034)	(0.733)	(57.754)	(0.080)	(0.070)	(0.574)	(0.105)
Total effect on treatment households	6.256***	0.404	-0.012	-0.708	-62.428	0.112	-0.060	-0.224	-0.029
(relative to control villages)	(0.380)	(0.784)	(0.043)	(1.072)	(63.744)	(0.107)	(0.080)	(0.630)	(0.135)
Mean of dep. var. in control villages	0.51	4.35	6.60	38.82	2329.06	4.32	-3.12	-9.32	0.15
S.D. of dep. var. in control villages	2.73	10.00	0.76	16.01	1112.02	1.47	1.29	9.84	2.08
R-squared	0.21	0.05	0.25	0.21	0.17	0.25	0.11	0.11	0.27
Observations	2836	2216	2216	2201	2215	2201	2193	2195	2123
Panel B: Round 3									
Control household in lean treatment village	1.510***	0.061	-0.005	0.088	-66.183	-0.008	-0.044	-0.635	-0.075
	(0.290)	(0.475)	(0.058)	(1.313)	(86.547)	(0.109)	(0.106)	(0.483)	(0.173)
Lean treatment household	3.719***	0.806*	0.012	-0.357	22.778	-0.024	0.180*	0.302	0.067
	(0.426)	(0.478)	(0.054)	(1.219)	(79.122)	(0.091)	(0.093)	(0.538)	(0.153)
Control household in harvest treatment village	-0.029	0.611	-0.051	-0.678	-98.414	-0.012	0.021	0.319	-0.101
	(0.164)	(0.575)	(0.046)	(1.172)	(72.154)	(0.105)	(0.110)	(0.507)	(0.158)
Harvest treatment household	0.251**	0.042	0.022	-0.433	16.980	-0.130	0.049	-0.597	-0.050
	(0.118)	(0.636)	(0.049)	(1.077)	(72.927)	(0.097)	(0.115)	(0.521)	(0.156)
Total effect on lean treatment households	5.230***	0.867	0.007	-0.269	-43.405	-0.032	0.136	-0.333	-0.008
(relative to control villages)	(0.384)	(0.546)	(0.045)	(1.131)	(69.163)	(0.104)	(0.092)	(0.460)	(0.154)
Total effect on harvest treatment households	0.223	0.652	-0.029	-1.110	-81.434	-0.143	0.070	-0.278	-0.151
(relative to control villages)	(0.168)	(0.492)	(0.047)	(1.116)	(69.102)	(0.104)	(0.093)	(0.476)	(0.153)
p-value: equal effects on lean treatment HHs and harvest treatment HHs	0.00	0.66	0.42	0.40	0.59	0.24	0.47	0.91	0.32
Mean of dep. var. in control villages	0.24	1.83	6.77	44.48	2631.72	4.82	-2.71	-5.69	0.14
S.D. of dep. var. in control villages	1.61	4.72	0.67	15.01	1153.57	1.37	1.33	7.13	1.98
R-squared	0.25	0.05	0.22	0.17	0.13	0.20	0.16	0.19	0.25
Observations	2799	2192	2197	2190	2197	2190	2175	2177	2138

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects. The omitted category is households in control villages
Panel A: The total effect on treatment households is the sum of the coefficients shown.
Panel B: The total effect on households treated with the lean season program is the sum of the coefficients on "control households in lean treatment village" and "lean treatment household" and the total effect on households treated with the harvest season program is the sum of the corresponding coefficients for the harvest season treatment.
See notes in Table 3 for dependent variable definitions.
All columns include baseline dependent variable, except for column 1, because MASAF labor data is not available.

Table 5: ITT Effects on Food Security and Labor Supply, By Region (Round 2)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Control household in treatment village in North	1.682	1.210	-0.215*	-4.870**	51.219	-0.251	-0.540**	-1.790*	-0.708**
	(1.250)	(1.044)	(0.128)	(2.132)	(214.226)	(0.197)	(0.166)	(0.925)	(0.273)
Treatment household in North	4.515***	0.738	0.050	-0.542	-456.952**	0.028	0.712***	2.077**	0.299
	(0.969)	(0.652)	(0.088)	(1.720)	(136.659)	(0.193)	(0.181)	(0.888)	(0.259)
Control household in treatment village in Central	0.986**	-1.322	-0.032	-1.190	-151.959	-0.045	-0.207	0.492	-0.144
	(0.456)	(1.669)	(0.077)	(1.892)	(108.772)	(0.214)	(0.151)	(1.055)	(0.265)
Treatment household in Central	3.895***	1.408	0.037	0.426	140.130*	0.163	0.080	-0.714	0.119
	(0.582)	(0.959)	(0.055)	(1.403)	(80.659)	(0.154)	(0.133)	(0.847)	(0.201)
Control household in treatment village in South	1.992***	0.337	0.032	0.133	15.713	0.239*	-0.124	-0.064	0.095
	(0.393)	(0.878)	(0.068)	(1.510)	(100.528)	(0.138)	(0.109)	(1.038)	(0.195)
Treatment household in South	5.187***	-0.230	-0.014	0.546	-16.516	-0.040	0.043	-0.351	-0.015
	(0.453)	(0.588)	(0.048)	(0.953)	(86.908)	(0.099)	(0.077)	(0.935)	(0.134)
Total effect on treatment households	6.196***	1.948*	-0.165	-5.411**	-405.733**	-0.223	0.172	0.287	-0.409
relative to control villages in North	(1.449)	(1.013)	(0.109)	(2.101)	(188.082)	(0.210)	(0.155)	(0.745)	(0.257)
Total effect on treatment households	4.881***	0.086	0.005	-0.764	-11.829	0.118	-0.127	-0.223	-0.025
relative to control villages in Central	(0.600)	(1.757)	(0.073)	(1.773)	(98.666)	(0.199)	(0.147)	(1.050)	(0.244)
Total effect on treatment households	7.179***	0.107	0.017	0.679	-0.803	0.198	-0.081	-0.415	0.080
relative to control villages in South	(0.409)	(0.810)	(0.062)	(1.497)	(95.802)	(0.136)	(0.109)	(1.004)	(0.188)
p-value: equal effects on control HHs in treatment villages in all regions	0.21	0.44	0.23	0.16	0.45	0.10	0.12	0.23	0.06
p-value: equal effects on treatment HHs in all regions	0.21	0.29	0.71	0.86	0.00	0.54	0.00	0.06	0.54
Mean of dep. var. in control villages (North)	1.67	1.67	7.38	52.97	3230.30	5.95	-2.66	-5.43	2.24
S.D. of dep. var. in control villages (North)	4.67	3.98	0.54	16.10	1168.71	1.07	1.42	5.86	1.63
Mean of dep. var. in control villages (Central)	0.41	5.91	6.47	34.33	1997.38	3.95	-2.87	-7.75	-0.17
S.D. of dep. var. in control villages (Central)	2.59	11.44	0.72	14.38	901.31	1.53	1.35	9.14	2.10
Mean of dep. var. in control villages (South)	0.36	3.85	6.49	38.41	2342.21	4.16	-3.44	-11.56	-0.19
S.D. of dep. var. in control villages (South)	2.22	9.74	0.72	15.01	1111.38	1.21	1.13	10.65	1.83
R-squared	0.22	0.06	0.25	0.22	0.18	0.25	0.12	0.11	0.28
Observations	2836	2216	2216	2201	2215	2201	2193	2195	2123

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects. The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region.
The omitted category is households in control villages.
See notes in Table 3 for dependent variable definitions.
All columns include baseline dependent variable, except for column 1, because MASAF labor data is not available.

Table 6: ITT Effects on Food Security and Labor Supply, By Region (Round 3)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Control household in lean treatment village in North	1.333**	-0.072	-0.233	-4.160	-178.557	-0.522*	-0.382*	-1.295	-0.865*
	(0.671)	(0.770)	(0.142)	(4.110)	(272.809)	(0.276)	(0.222)	(1.232)	(0.459)
Lean treatment household in North	4.930***	0.933	0.285	7.140**	281.205	0.516**	0.545**	0.930	0.958**
	(1.049)	(0.693)	(0.183)	(2.437)	(292.515)	(0.159)	(0.241)	(0.915)	(0.435)
Control household in harvest treatment village in North	-0.227	0.366	-0.144	-4.269	-29.143	-0.557*	-0.540**	-2.295*	-0.953**
	(0.409)	(0.761)	(0.090)	(3.770)	(116.815)	(0.282)	(0.271)	(1.370)	(0.407)
Harvest treatment household in North	0.104	0.977*	0.164*	3.064*	-41.926	0.201*	0.783**	2.315*	0.788**
	(0.395)	(0.568)	(0.098)	(1.708)	(149.488)	(0.118)	(0.327)	(1.362)	(0.381)
Control household in lean treatment village in Central	0.973**	0.417	0.026	0.295	-45.950	0.137	0.179	0.347	0.164
	(0.475)	(0.864)	(0.098)	(2.265)	(142.586)	(0.174)	(0.178)	(0.554)	(0.287)
Lean treatment household in Central	3.199***	1.538	-0.023	0.697	-27.470	0.009	0.105	-0.066	0.015
	(0.632)	(1.021)	(0.076)	(2.053)	(109.431)	(0.153)	(0.156)	(0.667)	(0.231)
Control household in harvest treatment village in Central	-0.306	1.620	0.102	2.222	-39.547	0.373**	0.197	0.741	0.396
	(0.321)	(1.167)	(0.080)	(2.068)	(127.970)	(0.177)	(0.154)	(0.535)	(0.253)
Harvest treatment household in Central	0.388*	-0.821	-0.128*	-2.775	-112.825	-0.339*	-0.176	-1.250*	-0.480*
	(0.209)	(1.567)	(0.069)	(1.837)	(120.869)	(0.183)	(0.178)	(0.673)	(0.256)
Control household in lean treatment village in South	1.898***	-0.437	0.022	0.843	-42.317	0.002	-0.160	-1.217	-0.075
	(0.421)	(0.628)	(0.079)	(1.567)	(118.244)	(0.152)	(0.148)	(0.836)	(0.232)
Lean treatment household in South	3.790***	0.446	-0.027	-3.072**	-10.753	-0.186	0.145	0.387	-0.121
	(0.631)	(0.520)	(0.075)	(1.522)	(112.092)	(0.119)	(0.125)	(0.947)	(0.204)
Control household in harvest treatment village in South	0.190	-0.010	-0.135**	-1.735	-165.593	-0.148	0.053	0.858	-0.221
	(0.181)	(0.655)	(0.065)	(1.381)	(103.921)	(0.136)	(0.162)	(0.821)	(0.214)
Harvest treatment household in South	0.206	0.264	0.084	0.137	134.771	-0.084	-0.013	-1.030	0.001
	(0.147)	(0.501)	(0.076)	(1.574)	(105.471)	(0.131)	(0.144)	(0.775)	(0.198)
Total effect on lean treatment households	6.263***	0.861	0.052	2.980	102.648	-0.006	0.162	-0.365	0.093
relative to control villages in North	(0.666)	(0.592)	(0.127)	(3.771)	(142.712)	(0.308)	(0.201)	(1.070)	(0.456)
Total effect on harvest treatment households	-0.123	1.343**	0.020	-1.205	-71.069	-0.356	0.243	0.020	-0.165
relative to control villages in North	(0.380)	(0.636)	(0.117)	(3.700)	(171.923)	(0.271)	(0.188)	(1.032)	(0.396)
Total effect on lean treatment households	4.173***	1.955**	0.003	0.992	-73.420	0.146	0.284*	0.282	0.179
relative to control villages in Central	(0.699)	(0.940)	(0.075)	(1.902)	(118.030)	(0.152)	(0.151)	(0.567)	(0.237)
Total effect on harvest treatment households	0.082	0.799	-0.026	-0.553	-152.372	0.034	0.021	-0.508	-0.084
relative to control villages in Central	(0.332)	(0.944)	(0.078)	(2.088)	(114.620)	(0.170)	(0.144)	(0.519)	(0.257)
Total effect on lean treatment households	5.688***	0.009	-0.005	-2.228	-53.069	-0.184	-0.014	-0.830	-0.195
relative to control villages in South	(0.523)	(0.676)	(0.064)	(1.351)	(101.623)	(0.147)	(0.128)	(0.800)	(0.213)
Total effect on harvest treatment households	0.396**	0.254	-0.051	-1.598	-30.822	-0.231	0.040	-0.172	-0.220
relative to control villages in South	(0.185)	(0.631)	(0.067)	(1.220)	(99.879)	(0.146)	(0.137)	(0.830)	(0.206)
p-value: equal effects on control HHs in lean treatment villages in all regions	0.34	0.72	0.26	0.52	0.90	0.13	0.12	0.20	0.17
p-value: equal effects on lean treatment HHs in all regions	0.37	0.61	0.27	0.00	0.61	0.00	0.28	0.68	0.08
p-value: equal effects on control HHs in harvest treatment villages in all regions	0.31	0.48	0.05	0.19	0.62	0.01	0.06	0.11	0.02
p-value: equal effects on harvest treatment HHs in all regions	0.72	0.44	0.03	0.07	0.28	0.04	0.04	0.06	0.02
Mean of dep. var. in control villages (North)	0.55	1.67	7.35	54.20	3109.72	6.26	-2.29	-3.55	1.80
S.D. of dep. var. in control villages (North)	2.44	5.15	0.57	17.01	1263.74	0.95	1.32	4.85	1.75
Mean of dep. var. in control villages (Central)	0.38	1.82	6.72	43.69	2440.25	4.48	-2.33	-3.11	0.13
S.D. of dep. var. in control villages (Central)	1.99	4.31	0.65	14.73	1133.34	1.30	1.31	4.44	1.88
Mean of dep. var. in control villages (South)	0.08	1.88	6.66	42.46	2650.42	4.69	-3.12	-8.24	-0.31
S.D. of dep. var. in control villages (South)	0.97	4.91	0.64	13.65	1103.74	1.27	1.23	8.33	1.87
R-squared	0.25	0.06	0.23	0.18	0.13	0.21	0.17	0.19	0.26
Observations	2799	2192	2197	2190	2197	2190	2175	2177	2138

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on households treated with the lean season program is the sum of the coefficients on "control households in lean treatment villages" and "lean treatment households" for each respective region and the total effect on households treated with the harvest season program is the sum of the corresponding coefficients for the harvest season treatment for each respective region.
The omitted category is households in control villages.
See notes in Table 3 for dependent variable definitions.
All columns include baseline dependent variable, except for column 1, because MASAF labor data is not available.

Table 7: ITT Effects on Durable Goods and Agricultural Inputs (Round 2)

Dependent variable:	(1) Ln (exp. fertilizer, 1st apply)	(2) Ln (exp. fertilizer, 2nd apply)	(3) Ln (kg fertilizer, 1st apply)	(4) Ln (kg fertilizer, 2nd apply)	(5) Indicator: used any fertilizer	(6) Indicator: owns any durables	(7) # of durables owned	(8) PCA index for durables
Control household in treatment village	0.010	0.087	0.145	0.090	0.027	-0.028	-0.167	-0.058
	(0.250)	(0.271)	(0.155)	(0.168)	(0.034)	(0.027)	(0.141)	(0.119)
Treatment household	0.047	0.147	0.041	0.146	0.008	0.021	0.127	-0.020
	(0.154)	(0.178)	(0.089)	(0.122)	(0.019)	(0.022)	(0.116)	(0.097)
Total effect on treatment households	0.057	0.235	0.186	0.236	0.035	-0.007	-0.040	-0.078
(relative to control villages)	(0.238)	(0.246)	(0.151)	(0.152)	(0.033)	(0.025)	(0.118)	(0.096)
Mean of dep. var. in control villages	5.26	3.41	2.95	1.85	0.65	0.82	2.89	0.09
S.D. of dep. var. in control villages	4.12	4.12	2.28	2.25	0.48	0.39	2.62	1.92
R-squared	0.23	0.11	0.32	0.16	0.30	0.10	0.42	0.45
Observations	2764	2764	2064	2064	2063	2213	2213	1520

OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
The PCA index for durable goods is only defined for households that own at least one good.
All columns include baseline dependent variable, except for column 1 and 2, because data on fertilizer expenditure is not available.

Table 8: ITT Effects on Durable Goods and Agricultural Inputs, By Region (Round 2)

Dependent variable:	(1) Ln (exp. fertilizer, 1st apply)	(2) Ln (exp. fertilizer, 2nd apply)	(3) Ln (kg fertilizer, 1st apply)	(4) Ln (kg fertilizer, 2nd apply)	(5) Indicator: used any fertilizer	(6) Indicator: owns any durables	(7) # of durables owned	(8) PC ind du
Control household in treatment village in North	2.193**	2.425***	0.911**	1.161**	0.223**	0.021	-0.068	0.0
	(0.836)	(0.625)	(0.423)	(0.366)	(0.103)	(0.036)	(0.276)	(0.
Treatment household in North	-0.442	-0.475	-0.283	-0.260	-0.062	-0.059**	-0.003	-0.
	(0.528)	(0.489)	(0.231)	(0.319)	(0.055)	(0.026)	(0.258)	(0.
Control household in treatment village in Central	-0.527	-0.318	0.030	0.155	-0.007	-0.008	-0.335	-0.
	(0.321)	(0.400)	(0.190)	(0.213)	(0.037)	(0.053)	(0.231)	(0.
Treatment household in Central	0.466	0.543	0.195	0.279	0.042	0.024	0.232	0.1
	(0.292)	(0.351)	(0.157)	(0.214)	(0.033)	(0.039)	(0.185)	(0.
Control household in treatment village in South	-0.097	-0.149	-0.016	-0.247	-0.007	-0.055	-0.041	0.1
	(0.328)	(0.358)	(0.256)	(0.268)	(0.055)	(0.035)	(0.199)	(0.
Treatment household in South	-0.114	0.035	0.006	0.137	-0.000	0.041	0.086	-0.
	(0.175)	(0.212)	(0.120)	(0.162)	(0.026)	(0.032)	(0.173)	(0.
Total effect on treatment households relative to control villages in North	1.751**	1.950***	0.628	0.901***	0.161	-0.038	-0.071	-0.
	(0.793)	(0.485)	(0.418)	(0.274)	(0.101)	(0.038)	(0.264)	(0.
Total effect on treatment households relative to control villages in Central	-0.061	0.225	0.224	0.434**	0.034	0.016	-0.104	-0.
	(0.328)	(0.373)	(0.184)	(0.204)	(0.032)	(0.049)	(0.222)	(0.
Total effect on treatment households relative to control villages in South	-0.211	-0.114	-0.010	-0.110	-0.008	-0.014	0.045	0.0
	(0.301)	(0.331)	(0.245)	(0.247)	(0.054)	(0.030)	(0.144)	(0.
p-value: equal effects on control HHs in treatment villages all three regions	0.01	0.00	0.14	0.01	0.10	0.31	0.60	0.1
p-value: equal effect on treatment HHs in all three regions	0.16	0.22	0.23	0.37	0.25	0.03	0.73	0.4
Mean of dep. var. in control villages (North)	3.19	2.37	1.90	1.44	0.42	0.97	4.10	0.4
S.D. of dep. var. in control villages (North)	4.15	3.70	2.29	2.11	0.50	0.18	2.42	1.9
Mean of dep. var. in control villages (Central)	6.88	4.24	3.67	2.00	0.78	0.79	2.95	0.3
S.D. of dep. var. in control villages (Central)	4.14	4.69	2.12	2.37	0.41	0.41	2.83	2.0
Mean of dep. var. in control villages (South)	4.51	3.02	2.68	1.84	0.61	0.80	2.51	-0.
S.D. of dep. var. in control villages (South)	3.70	3.66	2.24	2.19	0.49	0.40	2.40	1.7
R-squared	0.24	0.12	0.32	0.17	0.30	0.10	0.42	0.4
Observations	2764	2764	2064	2064	2063	2213	2213	15

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region. The omitted category is households in control villages.
The PCA index for durable goods is only defined for households that own at least one good.
All columns include baseline dependent variable, except for column 1 and 2, because data on fertilizer expenditure is not available.

Table 9: Price index (Round 2)

	(1) National	(2) North	(3) Center	(4) South
Treatment villages	-0.519	-5.542	-33.901	29.584
	(17.052)	(19.896)	(26.482)	(27.595)
Mean of dep. var. in control villages	485.20	554.04	512.18	452.54
S.D. of dep. var. in control villages	90.25	61.12	77.87	92.02
R-squared	0.28	0.63	0.10	0.27
Observations	179	22	62	95

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Observations are at EA level. Estimates include district and week-of-interview fixed effects.

Table 10: Test of Labor Market Tightening (Round 2)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Control household in treatment village (No hired labor)	1.653***	-0.646	-0.025	-1.380	-44.851	0.034	-0.171**	0.076	-0.160
	(0.383)	(0.973)	(0.050)	(1.127)	(79.023)	(0.112)	(0.086)	(0.685)	(0.146)
Treatment household (No hired labor)	5.448***	0.790	-0.002	0.023	-32.331	0.045	0.092	-0.324	0.045
	(0.424)	(0.585)	(0.038)	(0.778)	(68.143)	(0.086)	(0.070)	(0.657)	(0.114)
Control household in treatment village (Hired labor)	1.539***	0.819	-0.009	0.039	23.241	0.194	-0.259	-0.017	0.082
	(0.456)	(0.872)	(0.081)	(1.819)	(118.191)	(0.191)	(0.171)	(1.058)	(0.264)
Treatment household (Hired labor)	3.433***	-0.164	0.048	1.027	-14.561	0.022	0.291**	0.172	0.126
	(0.505)	(0.585)	(0.063)	(1.558)	(106.186)	(0.145)	(0.138)	(0.899)	(0.199)
Any hired labor in previous season	0.280	-2.824**	0.114*	1.566	229.647**	0.248*	0.273*	2.030**	0.328
	(0.336)	(1.062)	(0.066)	(1.315)	(99.388)	(0.146)	(0.157)	(0.844)	(0.218)
Total effect on treatment households	7.102***	0.144	-0.027	-1.357	-77.182	0.079	-0.079	-0.248	-0.115
relative to control villages (No hired labor)	(0.444)	(1.001)	(0.044)	(1.088)	(71.387)	(0.110)	(0.083)	(0.640)	(0.136)
Total effect on treatment households	4.972***	0.654	0.039	1.067	8.680	0.216	0.032	0.155	0.208
relative to control villages (Hired labor)	(0.505)	(0.784)	(0.076)	(1.709)	(107.416)	(0.167)	(0.158)	(1.026)	(0.238)
Mean of dep. var. in control villages	0.51	4.35	6.60	38.82	2329.06	4.32	-3.12	-9.32	0.15
S.D. of dep. var. in control villages	2.73	10.00	0.76	16.01	1112.02	1.47	1.29	9.84	2.08
R-squared	0.22	0.07	0.26	0.22	0.19	0.26	0.12	0.12	0.29
Observations	2836	2216	2216	2201	2215	2201	2193	2195	2123

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients shown, respectively for households without and with any hired labor in the previous season. The omitted category is households in control villages. See notes in Table 3 for dependent variable definitions.
All columns include baseline dependent variable, except for column 1, because MASAF labor data is not available.

Table 11: ITT Effects on Other Program Participation and Transfer/Loan (Round 2)

Dependent variable:	(1) Indicator: any free food or maize	(2) Indicator: any school feeding	(3) Indicator: any scholarship	(4) Indicator: any cash transfer	(5) Indicator any fertilizer coupon	(6) Indicator: any transfers received	(7) Indicator: any transfers given	(8) Indicator: any loans received	(9) Indicator: any loan payments
Control household in treatment village	0.012	0.013	-0.000	0.000	0.032	0.015	-0.048**	-0.029	0.002
	(0.030)	(0.032)	(0.005)	(0.004)	(0.031)	(0.026)	(0.020)	(0.023)	(0.012)
Treatment household	-0.019	-0.007	-0.003	0.004	0.018	-0.028	0.007	-0.002	0.001
	(0.015)	(0.021)	(0.004)	(0.005)	(0.021)	(0.020)	(0.016)	(0.016)	(0.011)
Total effect on treatment households	-0.006	0.005	-0.003	0.004	0.051*	-0.013	-0.041**	-0.031	0.003
(relative to control villages)	(0.029)	(0.031)	(0.004)	(0.005)	(0.029)	(0.025)	(0.019)	(0.022)	(0.011)
Mean of dep. var. in control villages	0.17	0.19	0.01	0.00	0.50	0.24	0.22	0.20	0.06
S.D. of dep. var. in control villages	0.37	0.40	0.08	0.07	0.50	0.42	0.41	0.40	0.25
R-squared	0.25	0.27	0.03	0.02	0.15	0.11	0.18	0.04	0.03
Observations	2219	2219	2219	2219	2840	2840	2840	2840	2840

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
Only column 1-4 include baseline dependent variable, because baseline data are not available for column 5-9.

References

Agarwal, Sumit, Vayacheslav Mikhed, and Barry Scholnick. 2016. "Does Inequality Cause Financial Distress? Evidence from Lottery Winners and Neighboring Bankruptcies." FRB of Philadelphia Working Paper No. 16-4.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2013. "Ordeal Mechanisms in Targeting: Theory and Evidence from a Field Experiment in Indonesia." National Bureau of Economic Research Working Paper 19127.

Angelucci, Manuela, and Giacomo DeGiorgi. 2009. "Indirect Effects of an Aid Program: How do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review*, 99(1): 486–508.

Bazzi, Samuel, Sudarno Sumatro, and Asep Suryahadi. 2015. "It's All in the Timing: Cash Transfers and Consumption Smoothing in a Developing Country." Working Paper.

Berhane, Guush, Daniel Gilligan, John Hoddinott, Neha Kumar, and Alemayehu Seyoum Taffesse. 2014. "Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Programme." *Economic Development and Cultural Change*, 63(1): 1–26.

Besley, T., and S. Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review*, 82(1): 249–261.

Blattman, Christopher, and Laura Ralston. 2015. "Generating Employment in Poor and and Fragile States: Evidence from Labor Markets and Entrepreneurship Programs." Working Paper.

Chen, Shaohua, Ren Mu, and Martin Ravallion. 2009. "Are there lasting impacts of aid to poor areas?" *Journal of Public Economics*, 93: 512–528.

Collins, Daryl, Jonathan Morduch, Stuart Rutherford, and Orlanda Ruthven. 2009. *Portfolios of the Poor: How the World's Poor Live on $2 a Day.* Princeton University Press.

Cunha, Jesse, Giacomo DeGiorgi, and Seema Jayachandran. 2014. "The Price Effects of In-Kind Transfers: Evidence from a Mexican Food Assistance Program." Working Paper.

Deiniger, Klaus, and Yanyan Liu. 2013. "Welfare and Poverty Impacts of India's National Rural Employment Guarantee Scheme." World Bank Policy Research Paper no. 6543.

Dillon, Brian, Peter Brummund, and Germano Mwabu. 2015. "How Complete are Labor Markets in Sub-Saharan Africa? Evidence from Panel Data in Four Countries." Working Paper.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2008. "How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya." *American Economic Review*, 98(2): 482–488.

Dutta, Puja, Rinku Murgai, Martin Ravallion, and Dominique vandeWalle. 2014. "Right to Work? Assessing India's Employment Guarantee Scheme in Bihar." World Bank Working Paper, https://openknowledge.worldbank.org/handle/10986/17195.

Frankenberg, Elisabeth, Jessica Y. Ho, and Duncan Thomas. 2016. "Oxford Handbook of Economics and Human Biology." , ed. John Komlos and Inas Rashad Kelly, Chapter Biological Health Risks and Economic Development. Oxford University Press.

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

Goldberg, Jessica. 2016. "Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi." *American Economic Journal: Applied Economics*, 7(1): 129–149.

Grosh, Margaret, Carlo del Ninno, Emil Tesliuc, and Azedine Ouerghi. 2008. *For Protection and Promotion: The Design and Implementation of Effective Safety Nets.* Washington DC:The World Bank.

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *Quarterly Journal of Economics*, Forthcoming.

Hoddinott, John, Guush Berhane, Daniel Gilligan, Neha Kumar, and Alemayehu Seyoum Taffesse. 2012. "The Impact of Ethiopia's Productive Safety Net Programme and Related Transfers on Agricultural Productivity." *Journal of African Economies.*

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

Jensen, Robert T., and Nolan H. Miller. 2011. "Do Consumer Price Subsidies Really Improve Nutrition?" *The Review of Economics and Statistics*, 93(4): 1205–1223.

Machinjili, Charles, and Shelton Kanyanda. 2012. "Integrated Household Survey 2010-2011: Household Socio-Economic Characteristics Report." Malawi National Statistical Office, Zomba, Malawi.

McKenzie, David. 2012. "Beyond Baseline and Follow-Up: The Case for More T in Experiments." *Journal of Development Economics*, 99(2): 210–221.

Miller, C. M., M. Tsoka, and K. Reichert. 2011. "The Impact of the Social Cash Transfer Scheme on Food Security in Malawi." *Food Policy*, 36(2): 230–238.

Murgai, Rinku, Martin Ravallion, and Dominique vandeWalle. Forthcoming. "Is Workfare Cost Effective against Poverty in a Poor Labor-Surplus Economy?" *World Bank Economic Review*.

National Statistical Office. 2015. "Statistical Yearbook." Government of Malawi.

Ravallion, Martin. 1999. "Appraising Workfare." *The World Bank Research Observer*, 14(1): 31–48.

Ravallion, Martin, Gaurav Datt, and Shubham Chaudhuri. 1993. "Does Maharashtra's Employment Guarantee Scheme Guarantee Employment? Effects of the 1988 Wage Increase." *Economic Development and Cultural Change*, 41(2): 251–275.

Ravi, Shamika, and Monika Engler. 2015. "Workfare as an Effective Way to Fight Poverty: The Case of India's NREGS." *World Development*, 67: 57–71.

World Bank. 2015. "The State of the Social Safety Nets 2015." Washington DC:World Bank.

Zimmermann, Laura. 2014. "Why Guarantee Employment? Evidence from a Large Indian Public-Works Program." University of Georgia Working Paper.

Zollmann, Julie. 2014. "Kenya Financial Diaries: Shilling Kwa Shilling - The Financial Lives of the Poor." FSD Kenya.

Zollmann, Julie, and Caitlin Sanford. 2016. "A Buck Short: What Financial Diaries Tell Us about Building Financial Services that Matter to Low-Income Women." Bankable Frontier Associates.

Online Appendix (not for publication)

Table A1: Balancing Test by Region (IHS3)

Dependent variable:	(1) Ganyu days	(2) Ln (p.c. food cons.) last week	(3) Food cons. score	(4) Per adult equivalent calories	(5) # food groups consumed	(6) Food security score	(7) Resilience index	(8) PCA index
Panel A: North								
Control household in treatment village	1.401	0.236**	-3.694	-112.693	-0.179	0.630	3.526	0.371
	(5.695)	(0.080)	(2.426)	(162.616)	(0.266)	(0.395)	(2.716)	(0.391)
Treatment household	-8.415	0.036	1.743	122.943	0.343**	-0.180	-1.055	0.100
	(8.012)	(0.084)	(2.299)	(120.019)	(0.164)	(0.164)	(0.676)	(0.168)
Total effect on treatment households	-7.014*	0.272***	-1.951	10.250	0.164	0.451	2.471	0.471
(relative to control villages)	(3.755)	(0.066)	(2.833)	(148.654)	(0.288)	(0.357)	(2.516)	(0.388)
Mean of dep. var. in control villages	20.40	6.34	45.15	2874.98	5.18	-2.35	-6.58	0.22
S.D. of dep. var. in control villages	36.64	0.51	12.92	1208.20	1.09	1.42	10.71	1.61
R-squared	0.16	0.11	0.20	0.14	0.19	0.18	0.23	0.18
Observations	318	318	316	318	316	318	318	315
Panel B: Central								
Control household in treatment village	17.717**	-0.004	-6.472***	26.011	-0.553**	-0.625***	-1.067	-0.721**
	(6.383)	(0.090)	(1.609)	(166.568)	(0.189)	(0.133)	(1.016)	(0.257)
Treatment household	-5.628	0.029	1.820	-38.583	0.132	0.196*	0.343	0.174
	(4.584)	(0.059)	(1.298)	(94.084)	(0.108)	(0.111)	(0.417)	(0.149)
Total effect on treatment households	12.089	0.026	-4.651***	-12.572	-0.421**	-0.429***	-0.724	-0.547**
(relative to control villages)	(7.626)	(0.080)	(1.497)	(163.164)	(0.192)	(0.122)	(0.982)	(0.238)
Mean of dep. var. in control villages	19.36	6.20	45.08	2512.70	4.98	-1.89	-3.26	0.22
S.D. of dep. var. in control villages	35.61	0.74	16.78	1058.91	1.37	1.34	7.85	2.03
R-squared	0.09	0.15	0.22	0.22	0.14	0.08	0.10	0.15
Observations	850	850	849	850	849	850	850	845
Panel C: South								
Control household in treatment village	7.238	-0.187**	-1.631	-164.096**	-0.273**	0.176*	0.533	-0.253
	(8.923)	(0.068)	(1.686)	(80.063)	(0.118)	(0.104)	(0.678)	(0.184)
Treatment household	1.802	0.044	0.280	52.909	0.123	0.132	0.200	0.158
	(7.229)	(0.049)	(1.072)	(78.775)	(0.083)	(0.085)	(0.474)	(0.123)
Total effect on treatment households	9.040	-0.143**	-1.351	-111.187	-0.150	0.308***	0.733	-0.095
(relative to control villages)	(7.900)	(0.062)	(1.563)	(73.279)	(0.111)	(0.101)	(0.659)	(0.179)
Mean of dep. var. in control villages	56.47	6.07	45.57	2482.46	5.12	-2.68	-5.03	-0.11
S.D. of dep. var. in control villages	93.15	0.70	15.20	1004.86	1.17	1.41	7.43	1.83
R-squared	0.17	0.20	0.15	0.12	0.21	0.35	0.27	0.27
Observations	1106	1106	1105	1106	1105	1106	1106	1096

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
Estimates include district and week-of-interview fixed effects.
See notes in Table 3 for dependent variable definitions.

Table A2: Balance of Household Characteristics (IHS survey)

Dependent variable:	(1) Household size	(2) Household head is female	(3) Household head attended 2nd school	(4) Indicator: owns any durables	(5) PCA index for durables	(6) Indicator: used any fertilizer	(7) ln (kg fertilizer, 1st apply)	(8) ln (kg fertilizer, 2nd apply)	(9) Indicator: any scholarship	(10) Indicator: any cash transfer
Control household in treatment village	0.148	-0.032	0.032	-0.005	0.066	-0.026	-0.169	0.228	0.000	0.010**
	(0.147)	(0.026)	(0.037)	(0.027)	(0.105)	(0.035)	(0.176)	(0.195)	(0.002)	(0.005)
Treatment household	-0.117	-0.013	0.000	0.024	-0.018	0.008	0.102	0.028	-0.001	-0.006
	(0.102)	(0.022)	(0.022)	(0.019)	(0.062)	(0.016)	(0.083)	(0.114)	(0.002)	(0.004)
Total effect on treatment households	0.031	-0.045*	0.032	0.019	0.048	-0.018	-0.067	0.256	-0.001	0.004
(relative to control villages)	(0.133)	(0.023)	(0.036)	(0.026)	(0.098)	(0.034)	(0.170)	(0.189)	(0.002)	(0.003)
Mean of dep. var. in control villages	4.67	0.26	0.21	0.82	-0.68	0.70	3.43	1.49	0.00	0.00
S.D. of dep. var. in control villages	2.31	0.44	0.40	0.39	1.27	0.46	2.30	2.28	0.05	0.07
R-squared	0.06	0.04	0.08	0.04	0.12	0.40	0.39	0.15	0.02	0.05
Observations	2274	2274	1651	2274	1850	2125	2126	2126	2274	2274

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
Estimates include district and week-of-interview fixed effects.
The PCA index for durable goods is only defined for HHs that own at least one good.

Table A3: ITT Effects on Food Security and Labor Supply, Full Sample without Baseline Control (Round 2 and 3)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: Round 2									
Control household in treatment village	1.589***	0.238	-0.059	-1.365	-82.728	-0.050	-0.156**	-0.215	-0.210
	(0.333)	(0.678)	(0.043)	(1.102)	(62.243)	(0.114)	(0.078)	(0.632)	(0.152)
Treatment household	4.667***	0.149	0.017	0.149	6.845	0.027	0.129**	0.146	0.098
	(0.338)	(0.420)	(0.031)	(0.669)	(50.781)	(0.068)	(0.061)	(0.491)	(0.097)
Total effect on treatment households	6.256***	0.387	-0.041	-1.216	-75.883	-0.023	-0.026	-0.069	-0.113
(relative to control villages)	(0.380)	(0.673)	(0.041)	(1.067)	(58.424)	(0.110)	(0.079)	(0.619)	(0.147)
Mean of dep. var. in control villages	0.51	4.43	6.58	38.30	2341.49	4.32	-3.12	-9.09	0.13
S.D. of dep. var. in control villages	2.73	9.71	0.74	15.66	1088.66	1.46	1.31	9.63	2.02
R-squared	0.21	0.04	0.17	0.14	0.13	0.18	0.08	0.09	0.16
Observations	2836	2836	2837	2825	2835	2825	2800	2802	2738
Panel B: Round 3									
Control household in lean treatment village	1.510***	-0.118	-0.091*	-0.795	-157.762**	-0.069	-0.060	-0.752*	-0.206
	(0.290)	(0.403)	(0.049)	(1.204)	(70.466)	(0.094)	(0.095)	(0.424)	(0.154)
Lean treatment household	3.719***	0.685	0.060	-0.446	111.186	0.011	0.159*	0.350	0.149
	(0.426)	(0.432)	(0.049)	(1.043)	(71.810)	(0.074)	(0.086)	(0.473)	(0.137)
Control household in harvest treatment village	-0.029	0.485	-0.113**	-2.060**	-124.704**	-0.135	0.027	0.049	-0.248*
	(0.164)	(0.473)	(0.040)	(1.044)	(58.383)	(0.090)	(0.090)	(0.420)	(0.136)
Harvest treatment household	0.251**	-0.053	0.028	-0.527	-1.224	-0.116	-0.029	-0.638	-0.076
	(0.118)	(0.504)	(0.041)	(0.885)	(62.316)	(0.083)	(0.097)	(0.443)	(0.133)
Total effect on lean treatment households	5.230***	0.567	-0.031	-1.241	-46.576	-0.059	0.099	-0.402	-0.057
(relative to control villages)	(0.384)	(0.471)	(0.044)	(1.143)	(63.518)	(0.093)	(0.086)	(0.396)	(0.143)
Total effect on harvest treatment households	0.223	0.432	-0.085**	-2.587**	-125.928**	-0.251***	-0.003	-0.589	-0.323**
(relative to control villages)	(0.168)	(0.420)	(0.043)	(1.065)	(59.127)	(0.091)	(0.083)	(0.413)	(0.136)
p-value: equal effects on lean treatment HHs and harvest treatment HHs	0.00	0.76	0.23	0.16	0.23	0.03	0.24	0.68	0.06
Mean of dep. var. in control villages	0.24	1.98	6.76	44.30	2658.77	4.78	-2.63	-5.15	0.16
S.D. of dep. var. in control villages	1.61	5.00	0.68	15.73	1142.64	1.35	1.36	6.88	1.97
R-squared	0.25	0.03	0.16	0.13	0.10	0.18	0.16	0.20	0.19
Observations	2799	2799	2804	2797	2804	2797	2774	2776	2746

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
The total effect on households treated with the lean season program is the sum of the coefficients on "control households in lean treatment village" and "lean treatment household" and the total effect on households treated with the harvest season program is the sum of the corresponding coefficients for the harvest season treatment. The omitted category is households in control villages.
See notes in Table 3 for dependent variable definitions.
No baseline outcome variable is included in regressions.

Table A4: ITT Effects on Food Security and Labor Supply, Baseline Subsample without Baseline Control (Round 2 and 3)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: Round 2									
Control household in treatment village	1.589***	0.032	-0.052	-1.590	-51.880	-0.031	-0.207**	-0.005	-0.193
	(0.333)	(0.789)	(0.051)	(1.240)	(70.856)	(0.126)	(0.085)	(0.688)	(0.170)
Treatment household	4.667***	0.457	0.028	0.692	-12.235	0.094	0.163**	-0.115	0.138
	(0.338)	(0.466)	(0.034)	(0.755)	(57.790)	(0.080)	(0.070)	(0.574)	(0.108)
Total effect on treatment households	6.256***	0.489	-0.024	-0.898	-64.115	0.064	-0.044	-0.120	-0.055
(relative to control villages)	(0.380)	(0.796)	(0.046)	(1.210)	(63.249)	(0.119)	(0.084)	(0.663)	(0.161)
Mean of dep. var. in control villages	0.51	4.35	6.60	38.82	2329.06	4.32	-3.12	-9.32	0.15
S.D. of dep. var. in control villages	2.73	10.00	0.76	16.01	1112.02	1.47	1.29	9.84	2.08
R-squared	0.21	0.05	0.19	0.16	0.15	0.21	0.10	0.10	0.18
Observations	2836	2216	2216	2201	2215	2201	2193	2195	2123
Panel B: Round 3									
Control household in lean treatment village	1.510***	-0.118	-0.091*	-0.795	-157.762**	-0.069	-0.060	-0.752*	-0.206
	(0.290)	(0.403)	(0.049)	(1.204)	(70.466)	(0.094)	(0.095)	(0.424)	(0.154)
Lean treatment households	3.719***	0.685	0.060	-0.446	111.186	0.011	0.159*	0.350	0.149
	(0.426)	(0.432)	(0.049)	(1.043)	(71.810)	(0.074)	(0.086)	(0.473)	(0.137)
Control household in harvest treatment village	-0.029	0.485	-0.113**	-2.060**	-124.704**	-0.135	0.027	0.049	-0.248*
	(0.164)	(0.473)	(0.040)	(1.044)	(58.383)	(0.090)	(0.090)	(0.420)	(0.136)
Harvest treatment household	0.251**	-0.053	0.028	-0.527	-1.224	-0.116	-0.029	-0.638	-0.076
	(0.118)	(0.504)	(0.041)	(0.885)	(62.316)	(0.083)	(0.097)	(0.443)	(0.133)
Total effect on lean treatment households	5.230***	0.567	-0.031	-1.241	-46.576	-0.059	0.099	-0.402	-0.057
(relative to control villages)	(0.384)	(0.471)	(0.044)	(1.143)	(63.518)	(0.093)	(0.086)	(0.396)	(0.143)
Total effect on harvest treatment households	0.223	0.432	-0.085**	-2.587**	-125.928**	-0.251***	-0.003	-0.589	-0.323**
(relative to control villages)	(0.168)	(0.420)	(0.043)	(1.065)	(59.127)	(0.091)	(0.083)	(0.413)	(0.136)
p-value: equal effects on lean treatment HHs and harvest treatment HHs	0.00	0.76	0.23	0.16	0.23	0.03	0.24	0.68	0.06
Mean of dep. var. in control villages	0.24	1.98	6.76	44.30	2658.77	4.78	-2.63	-5.15	0.16
S.D. of dep. var. in control villages	1.61	5.00	0.68	15.73	1142.64	1.35	1.36	6.88	1.97
R-squared	0.25	0.03	0.16	0.13	0.10	0.18	0.16	0.20	0.19
Observations	2799	2799	2804	2797	2804	2797	2774	2776	2746

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
The total effect on treatment households is the sum of the coefficients shown. The omitted category is households in control villages.
The total effect on households treated with the lean season program is the sum of the coefficients on "control households in lean treatment village" and "lean treatment household" and the total effect on households treated with the harvest season program is the sum of the corresponding coefficients for the harvest season treatment. The omitted category is households in control villages.
See notes in Table 3 for dependent variable definitions.
No baseline value of the outcome variable is included in regressions, but the sample is restricted to HHs whose baseline dependent variable is available.

Table A5: ITT Effects on Food Security and Labor Supply, By Region (Round 2, including additional HH characteristics)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Control household in treatment village in North	1.154	1.760	-0.285**	-5.616**	-209.679	-0.274	-0.590**	-1.903**	-0.853**
	(1.203)	(1.132)	(0.113)	(2.076)	(225.359)	(0.179)	(0.184)	(0.958)	(0.297)
Treatment household in North	4.848***	0.601	0.067	0.161	-373.891**	0.148	0.815***	2.641**	0.458*
	(1.336)	(0.763)	(0.091)	(1.908)	(130.821)	(0.173)	(0.191)	(0.833)	(0.237)
Control household in treatment village in Central	1.266*	-0.308	0.016	-0.948	-88.058	0.055	-0.091	1.024	0.017
	(0.647)	(1.535)	(0.088)	(1.962)	(116.640)	(0.229)	(0.162)	(1.032)	(0.293)
Treatment household in Central	3.978***	1.361	0.051	0.298	128.918	0.070	0.075	-0.742	0.082
	(0.756)	(1.065)	(0.065)	(1.558)	(97.474)	(0.164)	(0.156)	(0.899)	(0.218)
Control household in treatment village in South	1.765**	1.709**	0.079	0.308	83.998	0.174	-0.205	-0.799	0.138
	(0.627)	(0.740)	(0.070)	(1.626)	(108.976)	(0.139)	(0.149)	(1.273)	(0.216)
Treatment household in South	4.981***	-0.602	-0.012	1.146	-45.386	0.043	0.072	0.566	0.042
	(0.524)	(0.676)	(0.055)	(1.080)	(106.181)	(0.116)	(0.095)	(1.003)	(0.144)
Total effect on treatment households relative to control villages in North	6.002***	2.360**	-0.218**	-5.455***	-583.570***	-0.126	0.225	0.738	-0.396
	(1.619)	(1.150)	(0.085)	(1.961)	(200.723)	(0.163)	(0.151)	(0.673)	(0.250)
Total effect on treatment households relative to control villages in Central	5.244***	1.053	0.067	-0.650	40.860	0.125	-0.016	0.283	0.099
	(0.701)	(1.577)	(0.077)	(1.838)	(95.205)	(0.213)	(0.149)	(0.990)	(0.267)
Total effect on treatment households relative to control villages in South	6.747***	1.107*	0.067	1.453	38.613	0.217	-0.133	-0.233	0.180
	(0.570)	(0.660)	(0.065)	(1.637)	(104.921)	(0.135)	(0.156)	(1.264)	(0.213)
p-value: equal effects on control HHs in treatment villages in all regions	0.81	0.49	0.02	0.08	0.38	0.14	0.11	0.12	0.02
p-value: equal effects on treatment HHs in all regions	0.55	0.25	0.66	0.86	0.01	0.88	0.00	0.02	0.31
Mean of dep. var. in control villages (North)	1.92	1.38	7.38	54.01	3307.25	5.96	-2.57	-5.20	2.33
S.D. of dep. var. in control villages (North)	4.98	2.59	0.54	15.91	1148.95	1.03	1.43	5.83	1.57
Mean of dep. var. in control villages (Central)	0.55	5.01	6.45	35.07	2013.06	4.01	-2.86	-7.55	-0.14
S.D. of dep. var. in control villages (Central)	3.07	9.69	0.74	15.23	898.35	1.55	1.36	9.06	2.17
Mean of dep. var. in control villages (South)	0.46	2.88	6.50	38.94	2367.70	4.30	-3.32	-10.69	-0.07
S.D. of dep. var. in control villages (South)	2.70	6.37	0.69	14.18	1093.19	1.11	1.25	10.65	1.82
R-squared	0.20	0.08	0.30	0.24	0.25	0.27	0.15	0.12	0.31
Observations	1606	1606	1606	1598	1605	1598	1589	1590	1551

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$

The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region. The omitted category is households in control villages. Estimates include district and week-of-interview fixed effects.

See notes in Table 3 for dependent variable definitions.

All columns include baseline dependent variable, except for column 1, because MASAF labor data is not available. All columns control for female headed HH, education of the head of the HH, an indicator for whether the head of the HH completed secondary school, the size of the household, and the number of children in the HH. Village and household level treatment indicators and baseline outcome variable are interacted with indicators for each of the three regions.

Table A6: ITT Effects on Food Security and Labor Supply, Northern Region (Round 2)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: full sample, no baseline control									
Control household in treatment village	1.846	1.488	-0.156	-3.208	9.852	-0.199	-0.246	-0.130	-0.356
	(1.227)	(1.003)	(0.108)	(2.649)	(184.244)	(0.241)	(0.163)	(1.009)	(0.309)
Treatment household	4.691***	0.437	0.052	-1.891	-378.307**	0.040	0.535**	1.402	0.174
	(1.091)	(0.735)	(0.082)	(2.101)	(104.680)	(0.197)	(0.195)	(0.926)	(0.289)
Total effect on treatment households	6.537***	1.925*	-0.104	-5.099*	-368.455**	-0.159	0.289	1.272	-0.182
(relative to control villages)	(1.619)	(0.953)	(0.091)	(2.539)	(162.519)	(0.240)	(0.170)	(0.972)	(0.292)
Mean of dep. var. in control villages	1.67	1.63	7.35	52.39	3210.67	5.92	-2.70	-5.63	2.16
S.D. of dep. var. in control villages	4.67	3.92	0.55	16.16	1155.05	1.07	1.42	5.88	1.67
R-squared	0.13	0.02	0.10	0.06	0.15	0.10	0.05	0.08	0.16
Observations	338	338	339	338	339	338	336	336	333
Panel B: baseline subsample, no baseline control									
Control household in treatment village	1.846	1.172	-0.208	-5.528*	48.679	-0.297	-0.462**	-0.881	-0.584
	(1.227)	(1.062)	(0.123)	(2.845)	(198.523)	(0.273)	(0.187)	(1.126)	(0.367)
Treatment household	4.691***	0.641	0.074	0.203	-434.228**	0.138	0.689**	1.945**	0.361
	(1.091)	(0.659)	(0.096)	(1.771)	(143.753)	(0.186)	(0.194)	(0.906)	(0.284)
Total effect on treatment households	6.537***	1.813*	-0.135	-5.325*	-385.549**	-0.159	0.227	1.064	-0.223
(relative to control villages)	(1.619)	(1.054)	(0.098)	(2.876)	(157.613)	(0.267)	(0.178)	(1.025)	(0.344)
Mean of dep. var. in control villages	1.67	1.67	7.38	52.97	3230.30	5.95	-2.66	-5.43	2.24
S.D. of dep. var. in control villages	4.67	3.98	0.54	16.10	1168.71	1.07	1.42	5.86	1.63
R-squared	0.13	0.02	0.10	0.07	0.15	0.12	0.08	0.09	0.18
Observations	338	309	310	307	310	307	307	307	301
Panel C: baseline subsample, including baseline outcome									
Control household in treatment village	1.846	1.109	-0.202	-4.715*	65.751	-0.216	-0.540**	-1.656*	-0.662**
	(1.227)	(1.075)	(0.131)	(2.385)	(191.029)	(0.248)	(0.163)	(0.959)	(0.298)
Treatment household	4.691***	0.686	0.053	-0.519	-458.481**	0.039	0.709**	2.119**	0.309
	(1.091)	(0.679)	(0.090)	(1.773)	(142.061)	(0.195)	(0.190)	(0.906)	(0.264)
Total effect on treatment households	6.537***	1.795	-0.149	-5.233**	-392.730**	-0.177	0.169	0.462	-0.353
(relative to control villages)	(1.619)	(1.052)	(0.108)	(2.419)	(145.592)	(0.246)	(0.156)	(0.825)	(0.269)
Mean of dep. var. in control villages	1.67	1.67	7.38	52.97	3230.30	5.95	-2.66	-5.43	2.24
S.D. of dep. var. in control villages	4.67	3.98	0.54	16.10	1168.71	1.07	1.42	5.86	1.63
R-squared	0.13	0.03	0.21	0.17	0.17	0.19	0.09	0.12	0.28
Observations	338	309	310	307	310	307	307	307	301

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region. The omitted category is households in control villages. Estimates include district and week-of-interview fixed effects.
See notes in Table 3 for dependent variable definitions.

Table A7: ITT Effects on Food Security and Labor Supply, Central Region (Round 2)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: full sample, no baseline control									
Control household in treatment village	1.233**	-0.331	-0.086	-2.746	-164.348*	-0.278	-0.275*	-0.967	-0.500*
	(0.548)	(1.470)	(0.077)	(1.767)	(93.300)	(0.196)	(0.146)	(0.950)	(0.276)
Treatment household	3.886***	1.105	0.012	0.746	83.730	0.166	0.158	0.138	0.182
	(0.586)	(0.888)	(0.050)	(1.217)	(78.766)	(0.121)	(0.121)	(0.772)	(0.179)
Total effect on treatment households	5.118***	0.774	-0.074	-2.000	-80.618	-0.112	-0.117	-0.829	-0.318
(relative to control villages)	(0.643)	(1.587)	(0.075)	(1.673)	(82.038)	(0.191)	(0.150)	(0.991)	(0.265)
Mean of dep. var. in control villages	0.41	5.45	6.48	34.41	2037.26	3.98	-2.88	-7.39	-0.11
S.D. of dep. var. in control villages	2.59	11.07	0.72	14.72	910.22	1.53	1.36	8.59	2.08
R-squared	0.15	0.04	0.05	0.03	0.02	0.05	0.06	0.04	0.04
Observations	988	988	990	988	988	988	966	966	926
Panel B: baseline subsample, no baseline control									
Control household in treatment village	1.233**	-0.352	-0.067	-2.877	-191.744*	-0.298	-0.258	-0.069	-0.427
	(0.548)	(1.682)	(0.089)	(2.122)	(113.138)	(0.226)	(0.155)	(1.062)	(0.312)
Treatment household	3.886***	1.222	0.045	0.877	138.989*	0.200	0.104	-0.587	0.189
	(0.586)	(0.953)	(0.055)	(1.465)	(83.115)	(0.153)	(0.132)	(0.831)	(0.200)
Total effect on treatment households	5.118***	0.870	-0.023	-2.000	-52.755	-0.097	-0.155	-0.655	-0.238
(relative to control villages)	(0.643)	(1.789)	(0.084)	(1.971)	(100.748)	(0.212)	(0.152)	(1.077)	(0.291)
Mean of dep. var. in control villages	0.41	5.91	6.47	34.33	1997.38	3.95	-2.87	-7.75	-0.17
S.D. of dep. var. in control villages	2.59	11.44	0.72	14.38	901.31	1.53	1.35	9.14	2.10
R-squared	0.15	0.05	0.06	0.03	0.02	0.06	0.06	0.04	0.04
Observations	988	821	822	819	821	819	808	808	771
Panel C: baseline subsample, including baseline outcome									
Control household in treatment village	1.233**	-0.805	-0.050	-2.391	-192.658*	-0.197	-0.240	-0.061	-0.320
	(0.548)	(1.658)	(0.082)	(1.838)	(113.327)	(0.205)	(0.152)	(0.986)	(0.270)
Treatment household	3.886***	1.354	0.036	0.465	141.385*	0.169	0.088	-0.670	0.132
	(0.586)	(0.953)	(0.056)	(1.420)	(81.188)	(0.155)	(0.134)	(0.843)	(0.203)
Total effect on treatment households	5.118***	0.549	-0.014	-1.927	-51.273	-0.028	-0.151	-0.730	-0.189
(relative to control villages)	(0.643)	(1.790)	(0.077)	(1.693)	(102.229)	(0.191)	(0.150)	(0.992)	(0.246)
Mean of dep. var. in control villages	0.41	5.91	6.47	34.33	1997.38	3.95	-2.87	-7.75	-0.17
S.D. of dep. var. in control villages	2.59	11.44	0.72	14.38	901.31	1.53	1.35	9.14	2.10
R-squared	0.15	0.06	0.09	0.08	0.04	0.09	0.07	0.05	0.11
Observations	988	821	822	819	821	819	808	808	771

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region. The omitted category is households in control villages. Estimates include district and week-of-interview fixed effects.
See notes in Table 3 for dependent variable definitions.

Table A8: ITT Effects on Food Security and Labor Supply, Southern Region (Round 2)

Dependent variable:	(1) MASAF days	(2) Ganyu days	(3) Ln (p.c. food cons.) last week	(4) Food cons. score	(5) Per adult equivalent calories	(6) # food groups consumed	(7) Food security score	(8) Resilience index	(9) PCA index
Panel A: full sample, no baseline control									
Control household in treatment village	2.050***	0.505	-0.045	-0.526	-113.138	0.037	-0.042	0.276	-0.085
	(0.432)	(0.709)	(0.061)	(1.497)	(85.926)	(0.137)	(0.089)	(0.853)	(0.193)
Treatment household	5.187***	-0.596	0.014	0.200	46.921	-0.064	0.022	-0.089	0.033
	(0.454)	(0.503)	(0.045)	(0.860)	(74.237)	(0.089)	(0.068)	(0.744)	(0.126)
Total effect on treatment households	7.237***	-0.092	-0.031	-0.326	-66.217	-0.027	-0.020	0.186	-0.052
(relative to control villages)	(0.449)	(0.613)	(0.056)	(1.440)	(84.021)	(0.131)	(0.089)	(0.804)	(0.181)
Mean of dep. var. in control villages	0.36	4.25	6.49	38.22	2382.57	4.24	-3.36	-10.94	-0.10
S.D. of dep. var. in control villages	2.22	9.38	0.70	14.62	1093.90	1.26	1.19	10.48	1.81
R-squared	0.27	0.04	0.06	0.09	0.06	0.09	0.05	0.05	0.05
Observations	1510	1510	1508	1499	1508	1499	1498	1500	1479
Panel B: baseline subsample, no baseline control									
Control household in treatment village	2.050***	0.199	-0.011	0.106	-27.629	0.147	-0.067	0.401	0.018
	(0.432)	(0.855)	(0.073)	(1.601)	(95.240)	(0.140)	(0.104)	(0.976)	(0.216)
Treatment household	5.187***	-0.232	0.001	0.676	-4.113	0.004	0.058	-0.336	0.039
	(0.454)	(0.596)	(0.050)	(0.982)	(85.953)	(0.103)	(0.078)	(0.944)	(0.142)
Total effect on treatment households	7.237***	-0.033	-0.010	0.783	-31.743	0.151	-0.009	0.065	0.057
(relative to control villages)	(0.449)	(0.789)	(0.066)	(1.588)	(89.745)	(0.136)	(0.104)	(0.937)	(0.206)
Mean of dep. var. in control villages	0.36	3.85	6.49	38.41	2342.21	4.16	-3.44	-11.56	-0.19
S.D. of dep. var. in control villages	2.22	9.74	0.72	15.01	1111.38	1.21	1.13	10.65	1.83
R-squared	0.27	0.03	0.07	0.10	0.07	0.11	0.04	0.04	0.06
Observations	1510	1086	1084	1075	1084	1075	1078	1080	1051
Panel C: baseline subsample, including baseline outcome									
Control household in treatment village	2.050***	0.158	0.024	0.593	-17.628	0.247*	-0.079	0.456	0.137
	(0.432)	(0.824)	(0.070)	(1.542)	(98.997)	(0.128)	(0.099)	(0.970)	(0.193)
Treatment household	5.187***	-0.236	-0.016	0.542	-16.962	-0.040	0.043	-0.363	-0.018
	(0.454)	(0.592)	(0.048)	(0.965)	(87.095)	(0.099)	(0.077)	(0.938)	(0.135)
Total effect on treatment households	7.237***	-0.079	0.008	1.135	-34.590	0.207	-0.036	0.092	0.119
(relative to control villages)	(0.449)	(0.753)	(0.064)	(1.542)	(93.349)	(0.127)	(0.101)	(0.930)	(0.184)
Mean of dep. var. in control villages	0.36	3.85	6.49	38.41	2342.21	4.16	-3.44	-11.56	-0.19
S.D. of dep. var. in control villages	2.22	9.74	0.72	15.01	1111.38	1.21	1.13	10.65	1.83
R-squared	0.27	0.04	0.15	0.16	0.11	0.18	0.06	0.04	0.18
Observations	1510	1086	1084	1075	1084	1075	1078	1080	1051

Notes: OLS estimates. Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
The total effect on treatment households is the sum of the coefficients on "control household in treatment village" and "treatment household" for each respective region. The omitted category is households in control villages. Estimates include district and week-of-interview fixed effects.
See notes in Table 3 for dependent variable definitions.

Table A9: Program take-up by region

	Round 1	Round 2	Round 3
Panel A: National			
Control village	0.05	0.04	0.02
Control household in treatment village	0.08	0.19	0.08
Treatment household in treatment village	0.10	0.57	0.24
Panel B: North			
Control village	0.09	0.13	0.05
Control household in treatment village	0.06	0.24	0.10
Treatment household in treatment village	0.10	0.58	0.28
Panel C: Central			
Control village	0.04	0.03	0.04
Control household in treatment village	0.06	0.13	0.06
Treatment household in treatment village	0.05	0.43	0.21
Panel D: South			
Control village	0.06	0.03	0.01
Control household in treatment village	0.11	0.22	0.09
Treatment household in treatment village	0.13	0.65	0.26

Table A10: TOT Effects on Food Security and Labor Supply (Round 2)

Dependent variable:	(1) Ganyu days	(2) Ln (p.c. food cons.) last week	(3) Food cons. score	(4) Per adult equivalent calories	(5) # food groups consumed	(6) Food security score	(7) Resilience index	(8) PCA index
HH participated in MASAF	0.962	-0.128	-3.477	-206.251	-0.104	-0.233	-0.230	-0.419
in last 30 days	(1.908)	(0.117)	(3.065)	(161.149)	(0.312)	(0.215)	(1.696)	(0.417)
Mean of dep. var. in control villages	4.45	6.57	38.23	2332.12	4.31	-3.12	-9.11	0.13
S.D. of dep. var. in control villages	9.73	0.74	15.60	1085.07	1.45	1.31	9.65	2.01
Observations	2820	2820	2805	2820	2805	2781	2783	2724

Notes: 2SLS estimates. HH MASAF participation instrumented with the indicator fot treatment village. The first-stage coefficient of the instrument is 0.34 and the first stage F-statistic is 186.07.
Standard errors clustered at the EA level. $*p < 0.10, **p < 0.05, ***p < 0.001$
Estimates include district and week-of-interview fixed effects.
See notes in Table 3 for dependent variable definitions.