



# The effects of SPIR livelihoods and nutrition interventions on women's and men's well-being

## Evidence from the SPIR midline survey

Melissa Hidrobo, Harold Alderman, Daniel O. Gilligan, Jessica Leight, Heleene Tabet

---

### Introduction

Graduation model programs complement targeted cash or food transfers with multisectoral investments in asset building, income generation, and access to markets to provide a “big push” to promote sustained poverty alleviation. Graduation model programs are being tested in many countries following the publication of results from a six-country study of programs styled after BRAC's Targeting the Ultra-Poor (TUP) program in Ethiopia, Ghana, Honduras, India, Pakistan, and Peru.<sup>1</sup> These programs led to substantial improvements in household economic outcomes as well as in some measures of mental health related to stress. However, the BRAC-style programs showed no effects on women's empowerment.

The Ethiopia SPIR graduation approach differs from the BRAC programs in two important ways. First, SPIR includes substantial, integrated programming designed to improve nutrition and women's empowerment. Second, SPIR provides smaller resource transfers (the largest being a one-time poultry or cash grant), and instead focuses on linking PSNP clients to information and improved public services. This learning brief provides evidence from an ongoing impact evaluation looking at whether the Ethiopia SPIR graduation approach improves gender norms, women's empowerment, and their well-being. We find that the enhanced SPIR livelihood and nutrition model has improved women's wellbeing especially with respect to gender norms, decisionmaking, and depression.

The SPIR Development Food Security Activity (DFSA) in Ethiopia is a five-year program (2016–2021) supporting implementation of the PSNP4 as well as complementary livelihood, nutrition, gender, and natural resource management activities intended to strengthen the program and expand its impacts. Under funding from USAID's Office of Food for Peace and in close collaboration with the Government of Ethiopia, World Vision leads implementation of the SPIR DFSA in partnership with the Organization

for Rehabilitation and Development in Amhara (ORDA) and CARE. SPIR DFSA targets more than 500,000 PNSP clients in 15 of the most vulnerable *woredas* in the Amhara and Oromia regions of Ethiopia. SPIR DFSA also incorporates a substantial learning agenda intended to use evidence to improve the design of the DFSA, provide feedback to strengthen its delivery, and draw lessons for both local government and other national and international stakeholders. World Vision, ORDA, and CARE provide guidance on the implementation of the learning agenda. IFPRI leads the planning and execution of the learning agenda activities.

## Evaluation design and interventions

### *Interventions*

The impact evaluation evaluates combinations of four SPIR interventions described below; L and N correspond to the primary SPIR interventions around livelihoods and nutrition, respectively, while L\* and N\* represent enhanced versions of these interventions.

- Intervention L:** SPIR livelihood activities: starting Village Economic and Social Associations (VESAs), financial literacy training, agriculture and livestock value chain development, home gardening and forage production
- Intervention L\*:** SPIR livelihoods activities **plus:** (i) Social Analysis and Action (SAA) to improve women's access to markets, (ii) aspirations promotion activities in randomly selected *kebeles* (subdistricts), and (iii) targeted poultry or cash livelihood transfers
- Intervention N:** SPIR nutrition activities: Nutrition Behavior Change Communication (BCC); WASH activities
- Intervention N\*:** SPIR nutrition activities **plus:** (i) Timed and Targeted Counseling (TTC) (more intensive nutrition BCC), (ii) Community-based Participatory Nutrition Promotion (CPNP), (iii) male engagement in BCC, and (iv) Interpersonal Therapy in Groups (IPT-G) interventions for women and men screened for depression (provided after the midline survey)

All four interventions have the potential to improve women's empowerment and well-being. The VESA and BCC activities under L and N bring women and men together to provide them with knowledge and training over a wide range of topics, thereby improving social capital and catalyzing women's empowerment. The L\* and N\* interventions go a step further by directly addressing depression through the IPT-G and directly addressing gender norms and women's role in the household through the SAA activities which address constraints on women's role in intrahousehold decision-making, mobility, choice of livelihood activities, and access to markets that derive from cultural and social norms; targeting women for the poultry and cash transfers; and including men in the nutrition BCC and creating male engagement groups. Positive social interaction and social support as a result of the group interventions and reduced poverty-related stress as a result of higher incomes (through the poultry and cash grants or other livelihood activities) are also pathways through which SPIR may lead to reductions in maternal and paternal depression, and IPV.<sup>2,3</sup> At the time of the midline survey, however, neither the male engagement groups nor IPT-G groups had begun.

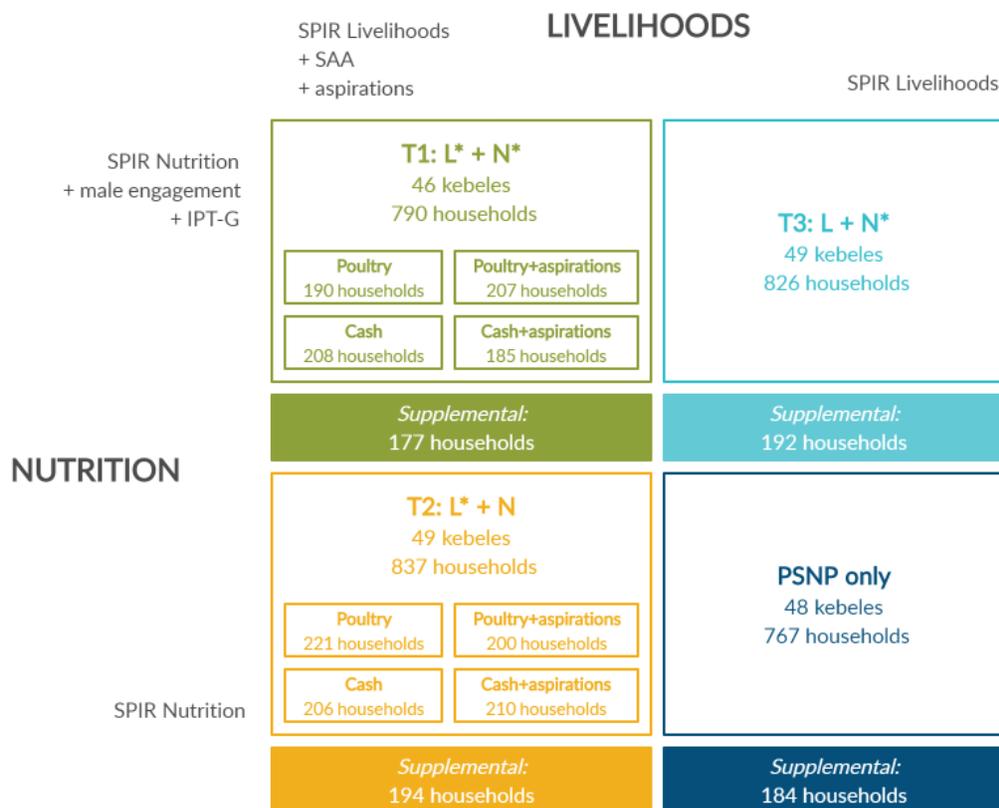
## Evaluation design

The impact evaluation uses a clustered randomized controlled trial (RCT) design to learn about the effect of different combinations of the SPIR intervention on the well-being of PSNP4 households: the livelihoods package (L), the nutrition package (N), and enhanced versions of each package (L\* and N\*, respectively). These packages were combined into multisectoral graduation model programs and randomized into four treatment arms: T1: L\*+N\*, T2: L\*+N, T3: L+N\*, T4: PSNP only. The evaluation sample comprises 192 *kebeles* in the Amhara and Oromia regions (Figure 1).

In addition, two supplemental interventions were cross-randomized across 94 *kebeles* in the experimental arms receiving the enhanced livelihoods intervention L\*: a one-time poultry package and a one-time cash transfer. Both interventions targeted poorer households (or more specifically, the poorest 10 out of the 18 households in each *kebele*). Half of the L\* *kebeles* were randomly selected to receive cash transfers targeted to women in these poorer households, and half of the L\* *kebeles* were randomly selected to receive poultry packages targeted to women in the poorer households.

Lastly, 50% of the L\* *kebeles* (n=47) were randomly assigned to receive an aspirations treatment (also described in more detail below). Randomization of the poultry/cash intervention and randomization of the aspirations treatment were balanced such that approximately 25% of L\* *kebeles* were assigned to either poultry only, poultry + aspirations, cash only, and cash + aspirations.

**Figure 1:** SPIR experimental impact evaluation design



## Midline survey

This brief draws on data from the midline survey conducted as part of the SPIR impact evaluation. All 3,314 households who met the inclusion criteria in the baseline sample were revisited between July and October 2019 in order to administer a survey to both the head of household and the primary female. Of these, 3,237 households were successfully surveyed, for an attrition rate of 2.3%. Our analysis here focuses primarily on gender norms, women’s decision-making, and measures of women’s well-being, including depression and intimate partner violence (IPV). We do not analyze the impact on women’s productive activities here, as this was presented in a separate learning brief.

## Analysis

In this brief, we present results in graphical form of the average treatment effect of each treatment (T1, T2, and T3) over all of its cross-randomized subtreatments (the “pooled” treatment effect of that treatment).<sup>4</sup> We focus on binary variables, and thus the graphs report the difference in probabilities of a certain outcome between households in either T1, T2, or T3 vis-à-vis households in the control arm. These differences are calculated conditional on *woreda* fixed effects. The x-axis corresponds to the effect size (that is, the difference in probabilities between T1, T2, or T3 households and control arm households) in percentage points. The graphs also report the 95% confidence interval corresponding to the coefficient estimated. If the 95% confidence interval excludes zero, then the effect is considered statistically significant: that is, we can reject the hypothesis that the observed difference between treatment and control households arises by chance.

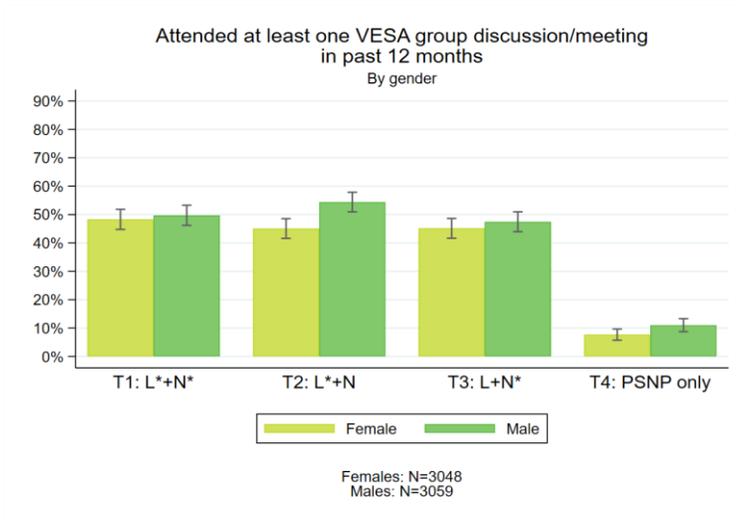
In addition to reporting the average treatment effect across the full sample, we report the average treatment effect among the poor who qualified for the poultry or cash sub-interventions (or more specifically, the poorest 10 out of 18 total households in each *kebele*) and nonpoor samples (the richest 8 out of 18 total households in each *kebele*). For the poor and nonpoor samples, we estimate the long form of the regression that includes indicators for each subtreatment when relevant (poultry, cash, aspirations), and then calculate the linear mean for the average treatment effect of T1, T2, and T3.

## Effects of the SPIR interventions on well-being

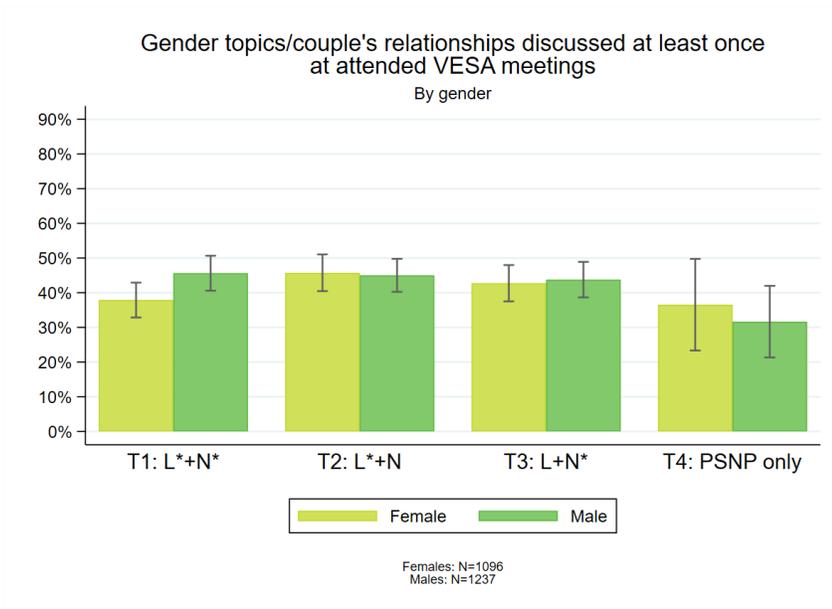
### *VESA attendance*

In order to interpret the effects of the interventions, we first report summary statistics from self-reported attendance in VESA group discussions, where most of the L activities took place. Figures 2a and 2b show the share of primary males and females reporting that they attended at least one VESA group discussion in the last 12 months, and among those who attended, the percentage that reported discussing gender topics at least once. Although the rates of VESA membership are high at the household level, at around 76–79% in T1, T2, and T3, the rates of attendance of the primary male or female are lower at around 50%. Among the primary males and females who attended a VESA meeting, a little less than 50% report discussing gender topics.

**Figure 2a: Attended at least one VESA meeting**



**Figure 2b: Discussed gender topics at VESA meeting**

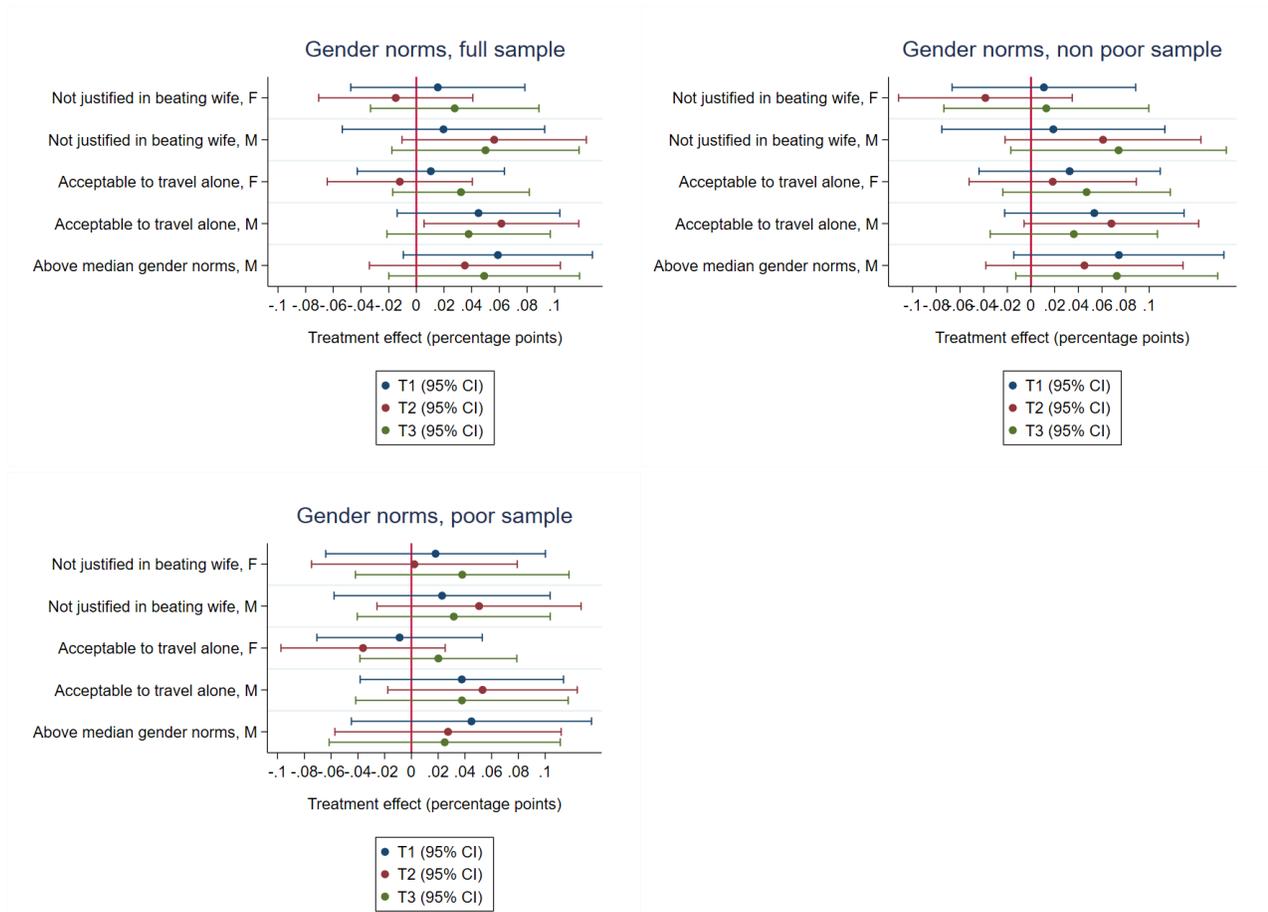


## *Gender norms*

We construct three different indicators of equitable gender norms. The first is from a list of four questions on whether a husband is justified in beating his wife under different circumstances. We create a binary indicator that equals one if the respondent says a husband is not justified in beating his wife under any of the four circumstances. The second is from a list of three questions on whether it is acceptable for a woman to travel alone to different places. We create a binary indicator that equals one if the respondent says it is acceptable for a woman to travel alone to the three specified places. Both these indicators on whether the husband is justified in beating his wife and whether it is acceptable for a woman to travel alone are created for the primary male and female. Lastly, for the primary male only, we ask him a list of 10 questions on gender inequitable norms where the response is strongly agrees (=0), somewhat agrees (=1), or does not agree (=2). Scores are summed up so that a higher score implies more equitable gender norms, and a binary indicator is created if scores are above the median.

Figure 3 reveals that SPIR improves men's but not women's attitudes/norms among the full sample. In particular, T2 increases the probability that men report that it is never justified for husbands to beat their wives (significant at the 10% level) and that it is acceptable for women to travel alone by approximately 6 percentage points each from a mean in the control group of 60% and 33% respectively. T1 also marginally increases the probability that a man is above the median on the gender equitable scale by 6 percentage points, from a mean in the control group of 44%. Among the poor sample, impacts of T1, T2, and T3 are not significant, but the poultry intervention does lead to marginal increases in the probability that a man reports it is always acceptable for a woman to travel alone and that he is above the median on the gender equitable scale. Among the nonpoor sample, T2 leads to marginal increases in the probability that a man reports it is acceptable for a woman to travel alone and T3 leads to marginal increases in the probability that a man reports it is not justified for a husband to beat his wife and that he is above the median on the gender equitable scale.

**Figure 3 Impacts on gender equitable norms**



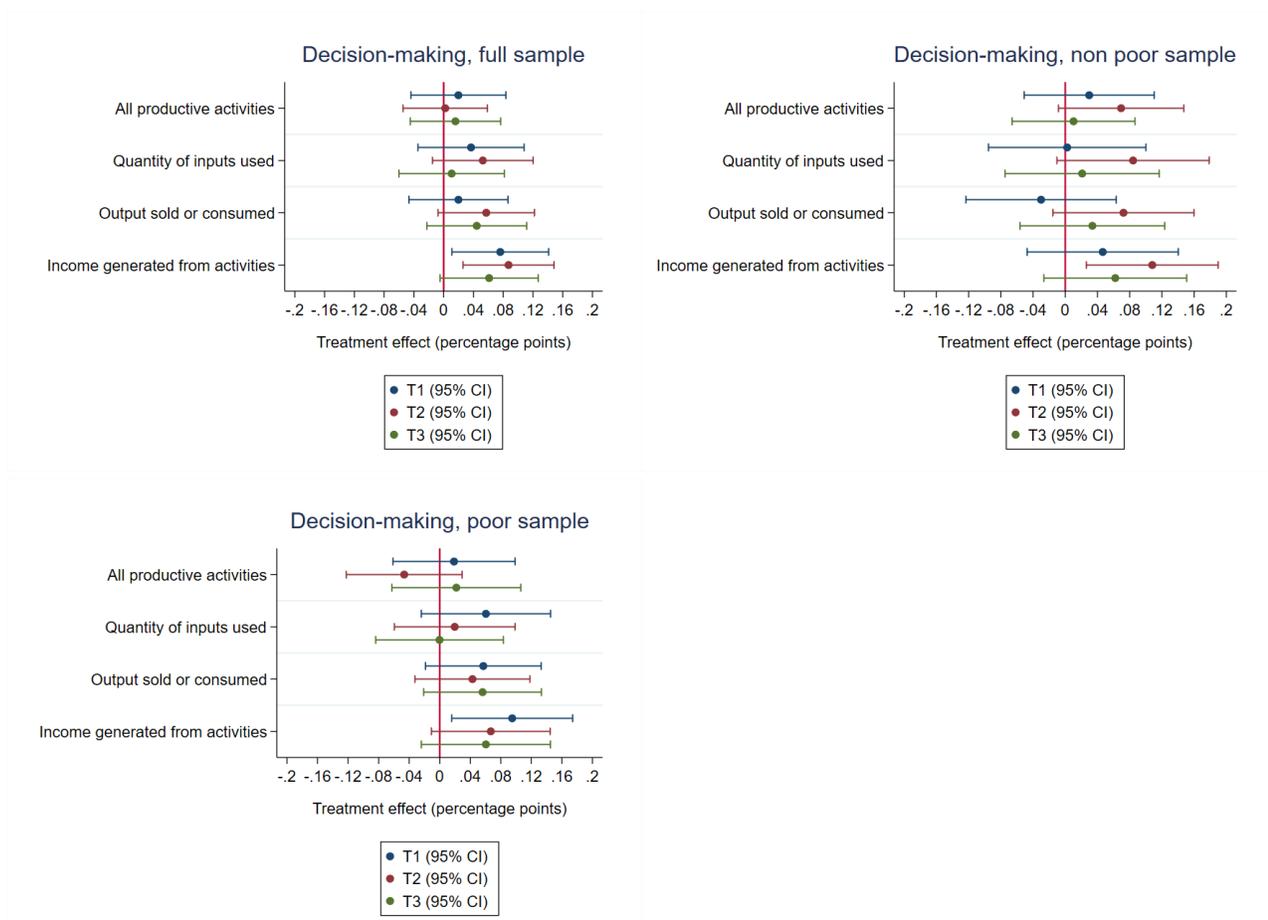
### Decision-making

We analyze impacts of SPIR on women’s decision-making with respect to input into productive decisions across 7 domains that include crop farming, horticulture, large livestock, small livestock, poultry, large household purchases and small household purchases. For each activity we ask 4 questions on decisionmaking. The first question is more general about how much input the respondent has in making decisions about the activity. The second through fourth questions – which are only asked about horticulture, large livestock, small livestock, and poultry - are more specific to the quantity of inputs that should be used; the quantity of the output that should be sold or consumed at home; and the use of income generated from the activity. We create 4 binary indicators that equal one if a woman reports having input into most decisions for all activities in which she participates. We use “input into most or all decisions” as opposed to “input into some decisions” because the share of women having input into some decisions is quite high (ranging from 84-92 percent), while the share of women having input into most decisions is much lower ranging from 40-55 percent.

Figure 4 reveals that among the full sample, all three arms lead to large increases in the probability of having input into most decisions with respect to the use of income generated from productive activities that ranges from 6.1-8.7 percentage points. T2 also leads to a marginally significant increase in decisionmaking on how much of the output of the activities should be consumed or sold. Among the subsample of nonpoor households, T1 and T3 have no impacts on women’s decisionmaking, but T2 in-

creates the probability that a woman has input into most decisions with respect to all productive activities, input use, and the use of income generated from productive activities. Impacts range from 6.9 to 10.8 percentage points and are significantly different from T1. Among the poor sample, T1 and T2 lead to large impacts on decisionmaking with respect to the use of income generated from the productive activities.

**Figure 4: Impact on women’s decision-making**



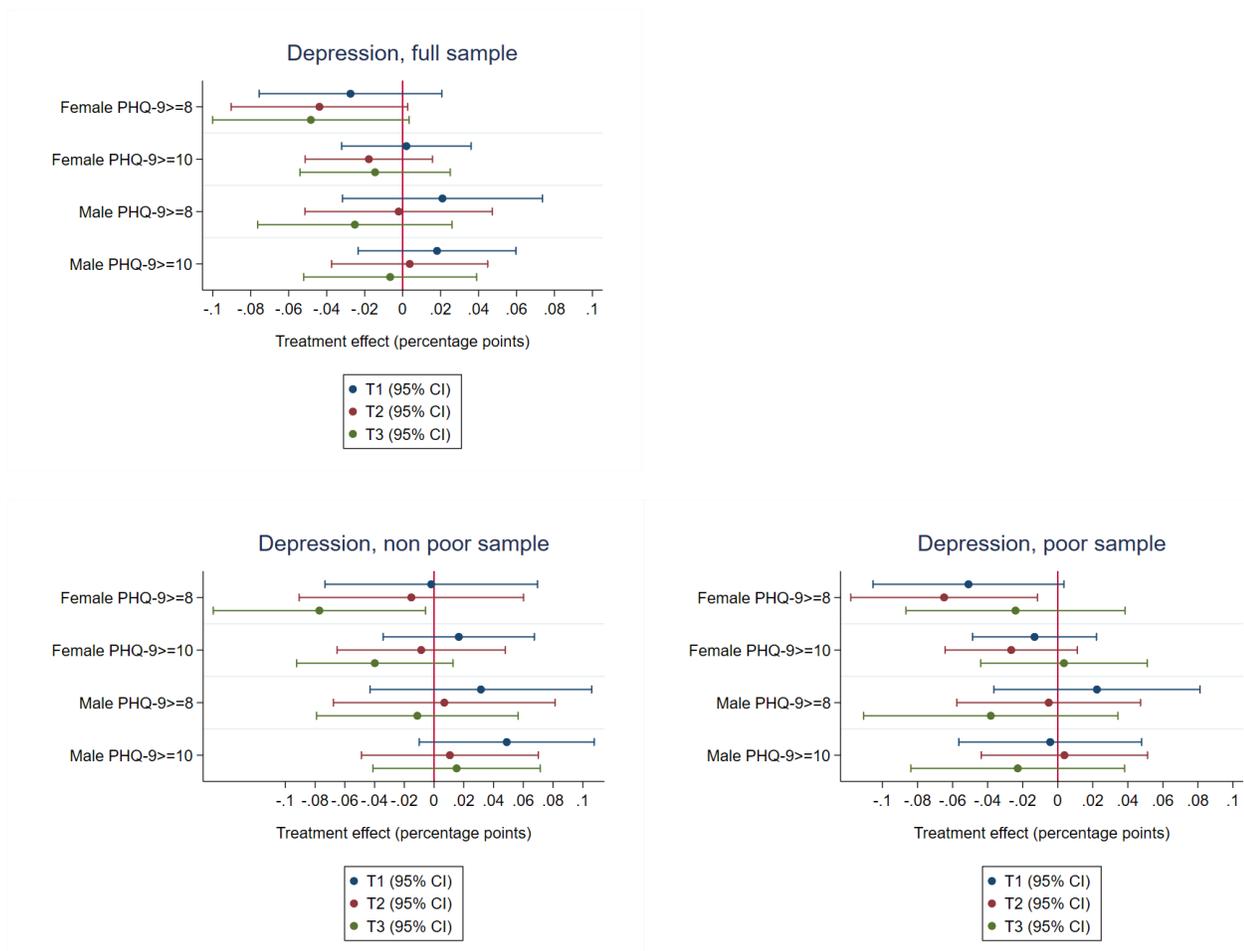
### Depression

The PHQ-9, a nine-item depression diagnostic instrument, was used to assess depressive symptoms of both the primary males and females.<sup>5,6</sup> The module asks respondents to indicate how frequently they experienced a set of depressive symptoms in the past two weeks, rating these on a scale of zero (never) to 3 (nearly every day). The PHQ-9 is then scored by adding up the nine responses, leading to a scale of zero to 27. There are various cutoffs for depression suggested in the literature. One that is commonly used defines mild depression as a reported PHQ-9 score between as 5 and 9, moderate depression between 10 and 14, moderately severe between 15 and 19, and with a score of 20 and above an individual is deemed to have symptoms of severe depression.<sup>7</sup> For recruitment into the IPT-G groups, we used a cutoff of 8 in order to increase the potential sample participating in the groups.

Figure 5 shows impact estimates for depressive symptoms using both binary indicators representing mild to severe depressive symptoms, using 8 as the cutoff, and moderate to severe depressive symptoms, using 10 as the cutoff. We present the results for both the primary male and female. Among the full sample, there are no impacts of T1, T2, or T3 on the probability of reporting a PHQ-9 score of 10 or higher for the primary male or female. However, there is a marginally significant impact from T2 and T3 on the probability that a woman has a PHQ-9 score of 8 or higher, implying that impacts are concentrated among women with relatively low levels of depressive symptoms. In particular, T2 and T3 decrease the probability of a woman having depressive symptoms at a score of 8 or higher by 4.4 to 4.8 percentage points, compared to a mean of 22% in the control group. Although these impacts are concentrated in T2 and T3, we cannot reject the hypothesis that the impacts are similar when compared with T1.

The impact of T2 on women’s depressive symptoms is mainly driven by poor households. Among poor households, the poultry intervention and T1 also lead to decreases in women’s depressive symptoms as measured by having a PHQ-9 score of 8 or higher. In contrast, the average impact of T3 on women’s depressive symptoms (PHQ-9 $\geq$ 8) is mainly from the nonpoor households. The impacts of T3 on women’s depressive symptoms are significantly different (at the 10 percent level) from T1 and T2, suggesting that for nonpoor households, the combination of L+N\* had significantly larger impacts than the other L and N combinations.

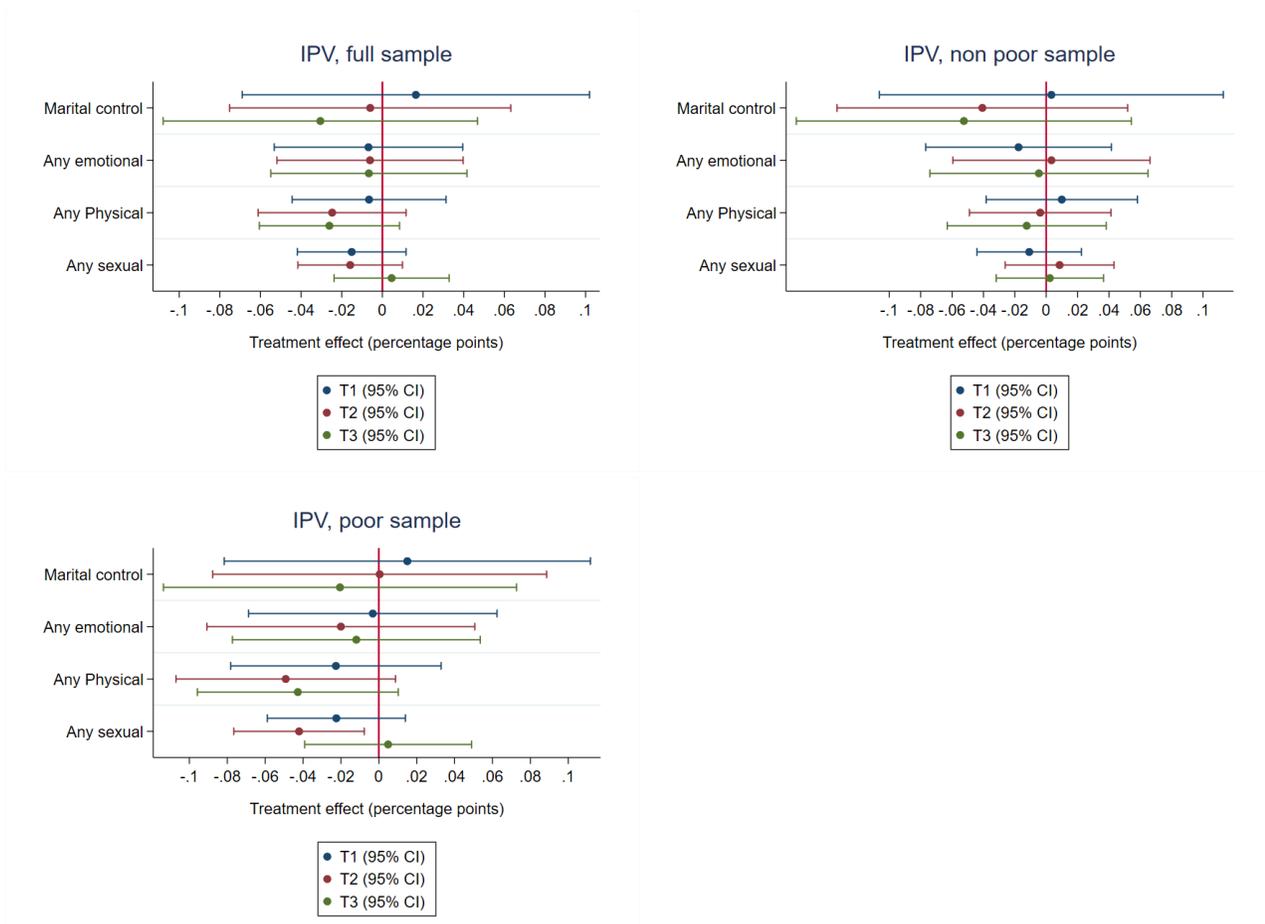
**Figure 5 Impacts on depressive symptoms**



## Intimate partner violence

Figure 6 reveals that there are no average treatment effects among the full sample of the SPIR intervention (T1, T2, T3) on marital control or emotional, physical, or sexual violence in the last 13 months. The coefficients are small in magnitude and statistically insignificant. Although there are no impacts on the full sample, in the poor subsample, T2 leads to marginally significant decreases in physical violence of 4.9 percentage points (or a 43% reduction compared to the control mean) and significant decreases in sexual violence of 4.2 percentage points (or a 61% reduction compared to the control mean). The effect on physical violence of T2 is mainly coming from the cash sub-intervention, while the effect on sexual violence of T2 is mainly coming from the poultry sub-intervention. Overall (across T1 and T2), poultry transfers lead to significant reductions in sexual violence, and cash transfers lead to marginally significant reductions in marital control.

**Figure 6 Impact on intimate partner violence**



## Conclusions

This brief presents evidence about the effectiveness of T1, T2, and T3 in improving gender norms, woman's decision-making, men's and women's depressive symptoms, and IPV. We find some evidence that T1 and T2 improve men's reported attitudes toward gender equitable norms among the full

sample. Both these arms included the L\* interventions and in particular the poultry and cash intervention, the SAA activities, and the aspirations videos, although we cannot reject the hypothesis that the impacts are similar when compared with T3. We find significant improvements in women's decisionmaking with respect to productive activities across all three arms among the full sample. We also find some evidence that T2 and T3 reduce the probability that a woman in our sample has a PHQ-9 score of 8 or higher, although we cannot reject the hypothesis that the impacts are similar when compared with T1. We find no impacts of any treatment arm on IPV for the full sample, but some large impacts of T2 on physical and sexual violence for the poor sample of households, and these impacts are significantly different from T3.

Given that not all activities had been rolled out by midline and some outcomes might take more time to change, such as personal norms regarding gender, the handful of positive impacts across T1, T2, and T3 are promising. The evidence on VESA group meeting attendance suggests that it may be possible to expand the impacts on gender norms and decisionmaking by encouraging even greater participation at VESA group meetings, where these issues are discussed. Taken together our results imply that the enhanced SPIR livelihood and nutrition model has improved women's wellbeing. Given that there are no significant differences in impact across treatment arms, we cannot speak to whether one particular intervention model was more effective than another at midline. Differences between models might appear however at endline.

---

---

## ACKNOWLEDGMENTS

We gratefully acknowledge funding for the impact evaluation from the United States Agency for International Development (USAID) under Cooperative Agreement No. AID-FFP-A-16-00008. This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM) led by the International Food Policy Research Institute (IFPRI). Additional funding support for this study was also provided by PIM. Michael Mulford, Chief of Party of the SPIR project, shares intellectual credit for study research design. We also thank our partners at World Vision, CARE, and ORDA for their contributions to the study and overall support. Thanks also to the research teams at Ambo University and Hawassa University for their feedback on the study design.

---

## ENDNOTES

<sup>1</sup> Banerjee A, Duflo E, Goldberg N, Karlan D, Osei R, Pariente W, Shapiro J, Thuysbaert B, and C.Udry. 2015. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236): 1260799-1-15.

<sup>2</sup> Lund, C., Brooke-Sumner, C., Baingana, F., Baron, E. C., Breuer, E., Chandra, P., ... Saxena, S. 2018. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. *The Lancet Psychiatry*, 5(4), 357–369. [https://doi.org/10.1016/S2215-0366\(18\)30060-9](https://doi.org/10.1016/S2215-0366(18)30060-9)

<sup>3</sup> Buller, A. M., Peterman, A., Ranganathan, M., Bleile, A., Hidrobo, M., & Heise, L. 2018. A mixed-method review of cash transfers and intimate partner violence in low- and middle-income countries. *World Bank Research Observer*

<sup>4</sup> See the SPIR midline report for a detailed description of the regression equations used to obtain the estimates reported in this brief.

<sup>5</sup> Bizu Gelaye, Michelle A. Williams, Seblewengel Lemma, Negussie Deyessa, Yonas Bahretibeb, Teshome Shibre, Dawit Wondimagegn, Asnake Lemenhe, Jesse R. Fann, Ann Vander Stoep, Xiao-Hua Andrew Zhou. Validity of the patient health questionnaire-9 for depression screening and diagnosis in East Africa, *Psychiatry Research*, Volume 210, Issue 2, 2013, Pages 653-661

<sup>6</sup> Woldetensay YK, Belachew T, Tesfaye M, Spielman K, Biesalski HK, Kantelhardt EJ, et al. (2018) Validation of the Patient Health Questionnaire (PHQ-9) as a screening tool for depression in pregnant women: Afaan Oromo version. *PLoS ONE* 13(2): e0191782. <https://doi.org/10.1371/journal.pone.0191782>

<sup>7</sup> Kroenke, K., Spitzer, R. L., & Williams, J. B. W. 2001. The PHQ-9. *Journal of General Internal Medicine*, 16, 606–613.

This publication was produced by IFPRI in collaboration with World Vision on the SPIR Program, Cooperative Agreement Number AID-FFP-A-16-00008, funded by the U.S. Agency for International Development.

This Learning Brief is made possible by generous support of the American people through the United States Agency for International Development (USAID). The contents are the responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

This publication has been prepared as an output of Strengthen PSNP4 Institutions and Resilience (SPIR) impact evaluation and has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by IFPRI.

## INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

*A world free of hunger and malnutrition*

IFPRI is a CGIAR Research Center

1201 Eye Street, NW, Washington, DC 20005 USA | T. +1-202-862-5600 | F. +1-202-862-5606 | Email: [ifpri@cgiar.org](mailto:ifpri@cgiar.org) | [www.ifpri.org](http://www.ifpri.org) | [www.ifpri.info](http://www.ifpri.info)

© 2020 International Food Policy Research Institute (IFPRI). This publication is licensed for use under a Creative Commons Attribution 4.0 International License (CC BY 4.0). To view this license, visit <https://creativecommons.org/licenses/by/4.0>.