Graduation Programs Targeting Women: Evidence from the Democratic Republic of Congo

Manuela Angelucci*, Rachel Heath†, Eva Noble‡

April 26, 2022§

Abstract

We study the impact of a graduation program for ultra-poor women in eastern Democratic Republic of Congo, a war-torn area. We cross-randomized the primary treatment–a holistic set of services and financial support–with a 16-week men's engagement program (MEP) for spouses and male household members. The core treatment has large effects on consumption, employment and finances, women's empowerment, and health, with most effects still significant two years after the program start and an internal rate of return of 19.9 percent. We find heterogeneous effects on intimate partner violence, which decreased for women at high risk of violence but increased for women at low risk. The MEP yields no lasting additional impacts. Multifaceted programs targeting women can be an effective way to lift people out of poverty and increase women's empowerment, although care is needed to minimize backlash.

Keywords: Poverty, graduation, women's empowerment, financial inclusion, fragile and conflict-affected areas

JEL classification: I15, I38, J22, O12

§We thank the participants of this study for sharing their time and experiences; the data collection firms Marakuja Kivu Research and Innovative Hub for Research in Africa (IHfRA), led by Adama Kabore and Aimable Amani, for their skill, local expertise, flexibility, and creativity to make this research successful; the Women for Women International program implementation staff: Abdoulaye Toure, Aloys Mateba, Thomas Wilondja, Ghislaine Maombi, Yvette Mudumbi, Gedeon Bamuleke, Chantal Kizungu, and Anna Mecagni; Liliana Ascencio, Frederic Cochinard, Olalekan Bello, Hanna Han, Samantha Archie, and Regina Villasor for their research assistantship; and Markus Goldstein and the WfWI Research Advisory Committee for their feedback. The project received ethical approvals by the RDC Ministere de la Santé Publique Comité National d'Éthique de la Santé (CNES) and University of Texas at Austin Institutional Review Board (2017-10-0148). This study is registered as Trial AEARCTR-0003437 in the AEA RCT Registry. This research was supported by the Ministry of Foreign Affairs of the Netherlands and BlackRock; the views expressed in this paper do not reflect the official views of the funders.

*Department of Economics, University of Texas at Austin, mangeluc@utexas.edu

†Department of Economics, University of Washington; rmheath@uw.edu

‡Monitoring, Evaluation, Research and Learning Department, Women for Women International, enoble@womenforwomen.org

1 Introduction

Graduation programs, which provide a holistic set of services and financial support to the ultra-poor, have shown that it is possible to make sustainable improvements in their lives by simultaneously addressing multiple barriers to poverty eradication (Banerjee et al., 2015). However, less is known about the impact of graduation programs that explicitly target women.

The effects of targeting women may differ from gender-neutral interventions. On one hand, targeting women may achieve more positive impacts for them and their households. For example, patriarchal norms and gender discrimination may prevent the flow of resources to women in untargeted interventions, potentially leading to inefficiencies. Additionally, if women are more altruistic, targeting them may maximize the program impacts on household consumption and well-being. Transferring resources to women may also increase their bargaining power (e.g., Attanasio and Lechene (2002)). On the other hand, targeting women poses a series of additional challenges. Discriminatory gender norms, from women's lack of financial autonomy to the normalization of violence against women (VAW), may pose additional barriers to women's socio-economic development. Backlash from partners could increase program beneficiaries' risk for intimate partner violence (IPV) (Angelucci, 2008; Buller et al., 2018; Heath et al., 2020), and defying social norms may worsen mental health.

This paper reports the impact of a graduation program targeting ultra-poor women in eastern Democratic Republic of Congo (DRC), a region facing protracted conflict in one of the poorest countries in the world. The intervention, implemented by Women for Women International (WfWI), pairs the graduation approach–providing capital, savings, enterprise development, mentoring, and social inclusion over 12 months–with training in women's rights, negotiation, decision-making, civic action, safety nets, and intentional building of social connections with other women. Since change in social norms cannot be achieved without men's involvement and agreement (Glinski et al., 2018), we cross-randomized the women's intervention with a men's engagement program (MEP), offered to spouses or male household members of half of the beneficiaries. The MEP consisted of 16 weekly discussion groups led by community leaders who had been previously trained on topics including VAW and women's economic empowerment, rights, and health.

Using a sample of 2039 women, we measure the program impacts one and two years after the program start on four families of outcomes: (i) household consumption and assets, and women's (ii) employment and finances, (iii) empowerment, and (iv) health. We find positive impacts on

all families. Specifically, non-durable consumption increases by 0.24SD and 0.17SD, livestock by 0.14SD and 0.24SD, and assets by 0.15SD and 0.09SD. The point estimates of these effect sizes are as large as or larger than the average effects from the six studies of bundled poverty alleviation programs reported in Banerjee et al. (2015), larger than the effects from a graduation intervention targeting ultra-poor women in Bangladesh (Bandiera et al., 2017) and smaller than the effects of a similar program in Afghanistan (Bedoya et al., 2019).

Comparing the consumption effect sizes of different interventions may be misleading, since costs may differ across program. Therefore, we consider the benefit-cost ratio (BCR) and internal rate of return (IRR) of different programs, assuming a five percent discount rate and that the program effects last in perpetuity.[1] Using only non-durable consumption, we estimate that the WfWI intervention breaks even after 4–5 years and has a BCR of 368 percent and an IRR of 19.9 percent. These are similar to two other graduation programs targeting women: the 20-year BCR and IRR in Bangladesh are 321 percent and 22.2 percent (Bandiera et al., 2017) and the 10-year BCR and IRR in Afghanistan are 232 percent and 26 percent, respectively. Conversely, the BCR and IRR from six graduation programs not targeting women appear to be lower: the BCRs range from -198 percent to 433 percent (the second largest one is 260 percent), while the IRRs range from an unspecified negative value to 23.4 percent (the second largest one is 13.5 percent) (Banerjee et al., 2015). This simple comparison suggests that graduation programs targeting women may increase consumption and assets as much as or more than gender-neutral interventions.

When we consider the other outcomes, we find that the average intent to treat (ITT) effect sizes one year since that start of the program are 0.08SD for employment and finances, 0.18SD for women's empowerment, and 0.04SD for health. These effects are as large at the end of the second year, showing that the improvements outlive the intervention and do not fade within this time interval. The effect sizes on employment and finances are comparable to similar interventions, while the health impacts are smaller than in other studies.

The point estimates of the impacts on women's empowerment are larger than the ones in Banerjee et al. (2015), whose sizes range between 0.01SD and 0.04SD and are never statistically significant.[2] This comparison suggests that it is possible to implement programs that empower women and are equally or more cost-effective than gender-neutral ones. Moreover, our findings

[1] While not necessarily realistic, these assumptions enable us to do a cost-benefit comparison of multiple graduation programs implemented in different countries.

[2] We can reject the hypothesis that our empowerment index and the women's decision-making index from the pooled studies are identical at conventional significance levels.

show that it is possible to improve the lives of ultra-poor women in fragile settings, despite limited infrastructure and governance, and in settings in which women are marginalized and discriminated against.

Two other notable findings pertain to treated women's partners and intra-household dynamics. First, we find that, while the average treatment effect on IPV is not statistically significant, the program has heterogeneous effects on this outcome: the program reduces IPV prevalence for women at above-average IPV risk, but it increases IPV for some women at low IPV risk. This heterogeneity is not unexpected. If IPV is a normal good for partners, women's empowerment can increase their bargaining power, thus reducing equilibrium IPV. At the same time, women's higher wealth and income may increase the returns of instrumental IPV (i.e., partners' acts of violence meant to gain control of spousal resources). Lastly, if a male partner's identity feels threatened by the woman's heightened status, he may retaliate by increasing IPV to assert his dominance. Therefore, the overall treatment effects on IPV can vary for different couples. Our findings highlight the importance of considering the distributional impacts of empowerment programs on IPV and including intervention elements to monitor and safeguard against instrumental violence.

Second, we find that the intervention generates large positive spillover effects within the household: child's attendance in school increases by 5 percentage points, and partners' income increases by 62 percent (whereas women's income increases by 19 percent). We explore the relationship between the treatment effects on IPV and partners' income and find that partners' income increases in households that experience reductions in IPV. This correlation suggests that partners' increases in income may result from increased cooperation between spouses or from women transferring resources to their partners in exchange for lower violence.

The Men's Engagement Program did not have any persistent additional impacts on any of the outcomes we considered. The lack of findings suggests that this type of intervention is insufficient in a setting in which anti-women norms are entrenched and endemic.

We conclude with the following thoughts. First, this study demonstrates that graduation programs can improve the well-being of the ultra-poor in extremely vulnerable, war-torn settings in which both governance and infrastructure are limited or non-existent, consistent with Bedoya et al. (2019) and Gibbs et al. (2020). Second, we find that many program effects are as large or larger two years after the program start than right after its conclusion. This dynamic suggests that the program benefits outlive the duration of the intervention and underscores the importance of collecting follow-up data. Third, we find that it is possible to improve the socioeconomic status

(SES) of women even in societies with strong anti-women norms: targeting women with cash, skill-building, knowledge transfers, and social networks improved women's income, entrepreneurship, savings, bargaining power and decision making. Conversely, the intervention targeting partners has no additional effects on women across measured outcomes. Different interventions may be needed to effect lasting change in men's attitudes and behaviors in this setting.

2 Background, Intervention, and Sampling

2.1 Background

The Democratic Republic of Congo (DRC) has faced conflict and instability for over 25 years, exacerbating poverty and gender inequalities for its population of 89.6 million. Despite its abundance of valuable natural resources, DRC is one of the world's poorest, least-developed, and gender-inequitable countries, ranking 179th out of 189 on both the Human Development Index and Gender Inequality Index (Conceição, 2019).

Decades of conflict have claimed over 5.4 million lives and displaced millions more, primarily in the eastern provinces (Coghlan et al., 2009; Zeender and Rothing, 2010). In South Kivu, our study area, continued insecurity and limited governance have hindered progress, disrupted markets, and destroyed assets. Almost 90 percent of informal entrepreneurs live below the poverty line and the vast majority are women: women are 6.7 times more likely than men to be "survivalist" entrepreneurs (Adoho and Doumbia, 2018).

A complex web of barriers reinforce women's vulnerability to poverty. Women often lack ownership of assets, personal income, and resources, and rarely receive inheritances. (Davis et al., 2014). 68 percent of women in South Kivu did not complete primary school, and only 60 percent of women aged 15 and older are literate, compared to 83 percent of men (MPSMRM et al., 2014). Women are treated as second-class citizens and face discrimination and abuse stemming from harmful gender norms that take precedence over national legal protections. Conflict exacerbates these abuses and adds others, including the prolific use of rape as a "weapon of war" and the normalization of sexual violence against women (VAW) among the civilian population (Bjørkhaug and Bøås, 2014). Almost half of the women in South Kivu have experienced physical violence and 35 percent have experienced sexual violence in their lifetime (MPSMRM et al., 2014).

2.2 Intervention

The key intervention is the “Stronger Nations Stronger Women" program implemented by WfWI, a bundled approach for ultra-poor women in conflict-affected settings that combines poverty alleviation and gender transformation objectives. This program aims to build women’s self-reliance in every aspect of life: economic stability, health and well-being, family and community participation, decision making, and social networks. The bundled approach supports social and economic empowerment through the following four activities:

1. Training on the value of women’s work, ways to save money, ways to earn income and improve income-generating activities, basic business skills, ways to improve health and well-being, women’s rights and prevention of VAW, strategies to make decisions and negotiate, civic action and advocacy, social networks, and safety nets;

2. Skill-building in numeracy and a chosen vocational skill (e.g., agriculture, sewing);

3. Resource provision in the form of a USD 10 monthly cash stipend (USD 120 total), formal and informal savings vehicles (e.g., village savings and loans associations (VSLAs), microfinance institutions), and referrals to health, legal, and financial services; and

4. Connections to other women through safe spaces for women to learn and share together, women-led social and economic groups, and a letter exchange with foreign supporters.

Over 12 months, participants were involved in two to five hours of programming per week, delivered to groups of 25 women in community-based training centers. This included biweekly social empowerment training sessions (24 sessions), weekly numeracy classes during months three and four (6 sessions), weekly business skills training during months four through 12 (30 sessions), and intensive vocational skills sessions for months seven through nine (up to 50 hours over 12 weeks). Participants also received training to set up their VSLAs, which then meet weekly.

In addition, half of the women who received the intervention were randomly assigned to the MEP study arm, meaning their spouses or other male household members participated in the MEP. Male community leaders, who were trained to share knowledge, facilitated 16 weekly discussion groups with the women’s male spouses, partners, or other household members on topics including women’s economic empowerment, domestic violence, women’s rights, and women’s health. Couples identified by WfWI staff members as high-risk for disputes or violence were also invited to join one couples’ dialogue session, a two-hour session with up to 25 couples on topics such as roles and responsibilities in a marriage, women’s rights (e.g., inheritance), civil registration,

and making commitments to reduce household conflict.

2.3 Sampling and Uptake

The WfWI field team screened and identified 2000 women as eligible for the program, with 39 additional eligible women screened to be replacements.[3] Specifically, members of the WfWI monitoring and evaluation team went to the pre-identified communities and explained the program and criteria to the local chiefs, who drew up lists of potential women. The women were then individually screened by the WfWI's M&E team. Standard eligibility criteria for participation in the program include: (1) experience with war/conflict (e.g., surviving violence, being displaced); (2) social vulnerability (e.g., poorer-than-average living conditions, facing restrictive traditional practices, or no or limited education); (3) economic vulnerability (e.g., extreme poverty, unemployment or limited to high risk or survivalist occupations); (4) motivation to participate in the full 12-month program; and (5) ability to participate (e.g., family support, adequate health). In this research study, an additional criterion was that women should be aged 18 to 55, and efforts were made to enroll only one household member to minimize spillover effects. The women were from the following communities in South Kivu: Ciheraoni-Luciga, Kamanyola, Mumosho, and Nyangezi.

We grouped the 2000 eligible participants into 80 clusters of 25 and randomly assigned clusters to control (C) and treatment groups (T) in equal proportions. Then, of the 40 clusters assigned to treatment, we cross-randomized 20 into the MEP group. The Stronger Women Stronger Nations program began August 2017 for 600 women and October 2017 for 400 women. The control group was scheduled to receive the intervention at a later date, after the end of our evaluation. We collected data at baseline (before the program start), endline (one year since the beginning of the intervention), and one year follow-up (two years since the beginning of the intervention).

2.4 Power Calculations

We found the minimum detectable standardized effect size considering a significance level of 0.05 and a power of 0.9. We assumed that the outcomes have a correlation coefficient of 0.5 between each of the three data waves and that the intra-cluster correlation is 0.1. Under these assumptions and using an ANCOVA estimator, we can detect a minimum impact size of 0.20SD for the main treatment and 0.27SD for the MEP treatment.

[3]No participant needed to be replaced. However, we include the 39 women in our sample in the analysis.

3 Outcomes

Our outcomes of interest fall into four broad domains: consumption, employment and finances, women's empowerment, and mental and physical health. For all indices, we normalize so that the control group has mean zero and standard deviation one at endline.

3.1 Consumption

We measure household consumption of 24 food items or groups in the past seven days, focusing on foods that are commonly consumed in the area.[4] We define food consumption as the total value of all the food items consumed by the household in the past seven days. We measure the quantity consumed (both purchased and non-purchased) and the unit values, which we use to create region-specific median prices. We then multiply the quantity consumed by the median price. We also measure an extensive set of non-food expenditures, which we convert to weekly values.[5] We winsorize consumption, income, earnings, and savings at the 5 percent level and convert them from Congolese francs (CDF) into US dollars (USD).[6]

We measure livestock ownership (cattle, goats, chickens, ducks, guinea pigs, and other livestock) by counting the number of animals in each category. Lastly, we count the number of durable assets owned in the household (chairs, beds, tables, fans, fridges, televisions, radios, water heaters, bed nets, lamps, buckets, machetes, baskets, watches, and hoes). We create livestock and asset indices, following Anderson (2008).

We consider the following specific aggregate outcomes: non-food expenditures, food consumption, livestock, and assets. Within non-food expenditures, we consider four sets of assignable goods: clothing and shoes for adult women, adult men, girls, and boys.

[4]The 24 food items are: cassava; sweet potatoes; potatoes; other tubers; rice; corn; other grains; legumes; plantains; eggs; milk; other dairy; onion; tomatoes; other vegetables; fruits; meat; fish; oil and ghee; tea and coffee; alcoholic and sugary drinks; dessert, sweets, and processed food; restaurant food; and spices and condiments.

[5]The items we measure are: communication; fuel; utilities; rent; housing; tobacco; toiletries and personal items; entertainment; medication; doctor/nurse fees; other medical expenses; religious festivals and activities; travel; transport; hotels; lottery tickets and gambling; firewood, kerosene, and charcoal (the previous month); fixing home damage, improvements, or expansions; charitable donations; weddings, funerals, dowries and bride prices; school/college fees; uniforms, books, or other supplies (the previous year); shoes and clothing for adult women, adult men, girls, and boys (the previous three months).

[6]We use exchange rates of 1600 per 1 USD for the baseline (2017) and endline (2018), and 1696 per 1 USD for the one year follow-up (2019).

3.2 Employment and Finances

We consider employment status, income, labor supply and time use, savings, and risk-taking behavior.

We consider a woman as working outside the home if she answers "yes" to the question "Did you work for pay, profit or gain during the past seven days?".[7] For respondents who said yes, we asked how many jobs they had, and then for up to two jobs, we measured job type and hours worked last week.[8] We measure self-employment from the respondent status in her primary occupation. To provide further information on how the treatment affected women's time allocation, we also conduct a time use module that measured market work, non-market work, childcare, leisure, and sleep yesterday.[9]

We consider both total earnings and earnings net of costs as labor outcomes. For each, we take the sum of earnings from primary and secondary jobs (when applicable) from the past seven days, including in-kind payments. For wage labor, we measure the period over which earnings are paid and how much was earned in that pay period and then calculate the implied earnings received over the past seven days.

We ask about business costs for both self-employed and wage-employed jobs, asking about both total spending over the past seven days and large tools or equipment required for one's business that required a large purchase. For these large purchases–defined as buying something they expected to use in their business for one month or more–we also ask the date of purchases (and deflate the cost to the time of the survey) and the expected number of months the item would be useful, then amortize its total cost to calculate the weekly use cost in current dollars. We also consider total costs as an outcome to indicate whether women are investing more in their businesses.

As a measure of the women's financial resources, we consider a woman's self-reported own monetary savings. We consider both a binary measure of whether she has any savings herself and a continuous measure of the amount of savings. We also consider whether the respondent reports being a member of a Village Savings and Loan Association (VSLA).

Finally, we measure risk-taking with a question that asked: "In general, are you someone

[7]To broadly capture all types of income-generating activities, the questionnaire further clarified that "This includes work done with payment in kind (such as for food), or helping out in the businesses of friends or family members, or selling goods and services. This includes work done in which you received no direct payment or compensation but friends or family members were compensated because of your work."

[8]Ninety-six percent of respondents who worked outside the home reported one or two jobs.

[9]The module captures multi-tasking by allowing up to one additional activity code for each primary activity.

who is willing to take risks or avoids taking risks?" A 5-item Likert scale measured the answers in ascending order of agreement.

3.3 Women's Empowerment

We group five outcomes related to women's empowerment: whether the woman is partnered (formally married or cohabiting with a partner); women's participation in household decisions; locus of control; intimate partner violence (IPV); and attitudes towards women.

To measure locus of control, we follow (Rotter, 1966). To measure women's participation in household decisions, we ask the respondents who makes the decision regarding the following household issues: whether the respondent can work outside the home, large household purchases, seeking medical care for the respondent, and seeking medical care for the respondents' children. We sort responses into three categories: (1) the respondent does not make the decision; (2) the respondent makes the decision with others; and (3) the respondent makes the decision alone. We award one point if the respondent made the decision alone or with others, then construct a decision-making index following (Anderson, 2008).[10]

We measure intimate partner violence (IPV) among women who are partnered at endline and one year follow-up.[11] Following the WHO's definition of IPV, we ask about physical, sexual, and emotional violence, and controlling behavior (World Health Organization et al., 2012).[12] We aggregate these occurrences into an index following Anderson (2008).

3.4 Mental and physical health

We use two self-reported scales to measure respondents' mental health–the Generalized Anxiety Disorder (GAD-7) Scale and the Patient Health Questionnaire (PHQ-9)–and a set of question about Activities of Daily Living (ADLs) to measure respondents' physical health. The GAD-7 Scale (Spitzer et al., 2006) assesses the respondent's level of generalized anxiety. The scale ranges from 0 to 21 and higher values correspond to higher anxiety levels. Scores of 5, 10 and 15 are cut-off points

[10]We also measured decision-making about the respondent's contraceptive usage. However, at baseline, 70 percent of respondents (and 67 percent of partnered respondents) said contraceptive usage was either "Not Applicable" to their household or that no decision was made concerning this issue. We accordingly exclude the contraceptive usage question from the decision-making index.

[11]Since answering these questions can be difficult or painful, we chose not to ask them at baseline.

[12]Specifically, we ask each respondent whether, in the past 12 months, their spouse or partner has: beaten or hit her; forced her to have sexual intercourse when she did not want to; insulted, yelled at, or threatened to hurt her; tried to prevent her from visiting or speaking to family or friends or seeking medical care; or taken her income, controlled how she spent money she earned, or gotten angry because of how she spent or saved her income.

for mild, moderate and severe anxiety. This instrument has been validated for the DRC (Mughal et al., 2020). The PHQ-9 is a screening tool for depression (Kroenke et al., 2001). Scores range from 0 to 27, where higher values correspond to more severe symptoms. Scores of 5–9 correspond to mild depression, scores of 10–14 correspond to moderate depression, scores of 15–19 correspond to moderately-severe depression, and scores of 20–27 correspond to severe depression.[13]

To measure ADLs, we ask the following questions: (a) "Can you currently do vigorous activities like running, lifting heavy objects, and carrying water?", (b) "Can you currently do moderate activities like working in the fields, sweeping, washing an infant, or walking 5 kilometers?", and (c) "How much physical pain have you experienced in the past month?" For each, we convert responses into binary variables where an answer of 1 means worse health, and construct an index out of these binary variables following Anderson (2008).

4 Identification and Estimation

We are interested in measuring the causal effect of the pooled treatments (T=1) on some outcome Y at endline (e) and follow-up (f) which are one and two years after the start of the intervention, respectively. To do so, we estimate the parameters of the following equation for participant i at time t:

$$Y_{it} = \beta_0 + \beta_1 T_i^e + \beta_2 T_i^f + \beta_3 f_t + X_i'\delta + \epsilon_i \tag{1}$$

The indicators T^e and T^f group treatment group members at endline and follow-up. The covariates X include the baseline value of Y (Y_{i0}), when available, a quadratic in woman's age, community fixed effects, and variables that are significantly different between treatment and control at baseline at the 10 percent level or higher (number of young children in household, hours of work, PHQ-9, and the pro-women attitudes index, as shown in Table 1 below).

The parameters β_1 and β_2 identify the intent to treat (ITT) effect at endline and one year follow-up.[14] These parameters are identified under random assignment and absent spillover effects. We believe that our design minimizes spillover effects, as the treated women and their husbands are

[13]The seventh question of the PHQ-9–"Trouble concentrating on things, such as reading the newspaper or watching television"–has a 9.4 percent non-response rate, so we drop it and report the score computed from the remaining eight questions. Therefore, the PHQ-9 levels we report slightly understate the severity of depression.

[14]Compliance was very high; 98 percent of respondents assigned to control did not receive treatment, and 94 percent of respondents assigned to treatment did receive treatment. Thus the ITT effects we present throughout are very close to treatment-on-the-treated estimates.

a very small fraction of the underlying community. We also test the hypothesis that the treatment effects are identical over time: $H_0 : \beta_1 = \beta_2$.

We estimate the parameters of all equations by OLS, clustering standard errors by group, for a total of 80 clusters. To account for multiple inference, we control for the false discovery rate within families of outcomes following Benjamini et al. (2006) and report the corresponding sharpened q-values in addition to the standard errors. Moreover, when applicable, we create indices for families of outcomes following Anderson (2008).

5 Baseline and Attrition

Table 1 shows mean outcomes at baseline for the treatment and control groups and the p-value of the test for equality between them. Out of 28 outcomes, four are statistically different between the control and treatment groups at the 90 percent significance level. This is slightly higher than what is expected to occur by chance, and a test of joint significance of the variables in Table 1 when regressed on treatment rejects the null of joint insignificance ($F(27, 1833) = 3.65$, $P < 0.001$).[15] However, these differences are generally small, and we do not find any clear patterns (e.g., control group women have slightly better health outcomes and slightly worse employment outcomes).

To address this issue, we add all the unbalanced baseline covariates to the set of controls of all our specifications, as mentioned above. We also re-estimate treatment effects using entropy weights to impose balance across arms on the covariates listed in Table 1 (Hainmueller, 2012; Hainmueller and Xu, 2013). The weighted and unweighted estimates are generally similar; see table A1. In the remainder of the paper, we follow our pre-analysis plan and present unweighted estimates.

Attrition was relatively low (7.5 percent at endline and 12 percent at one-year follow-up) and did not vary between treatment and control arms; see table A2. Participants in the control arm expected to be assigned to WfWI treatment at a later stage, which may have motivated them to continue to engage with the survey teams.

[15]Similarly, we find that not all variables are balanced if we separate the two treatment arms and compare each mean to the control arm.

Table 1: Balance test

Variable	Mean Value in Control	Mean Value in Treated	P-value diff in means	N
Children under 5 in house	1.394	1.266	0.049	2,039
Respondent age	31.149	31.438	0.708	2,036
Respondent literate	0.299	0.271	0.244	2,039
Non-food consumption (USD)	2.561	2.693	0.465	2,039
Food consumption (USD)	0.536	0.570	0.298	2,039
Women's clothes (USD)	0.095	0.094	0.894	2,039
Men's clothes (USD)	0.004	0.004	0.344	2,039
Girls' clothes (USD)	0.013	0.012	0.837	2,039
Boys' clothes (USD)	0.007	0.006	0.466	2,039
Assets (standardized)	-0.008	-0.048	0.540	2,039
Livestock (standardized)	0.013	-0.024	0.619	2,039
Total earnings (USD)	0.848	0.932	0.422	2,039
Earnings net of costs (USD)	0.532	0.658	0.152	2,031
Total business costs (USD)	0.263	0.213	0.174	2,031
Worked last week	0.414	0.415	0.984	2,039
Hours of work last week	7.459	10.425	0.007	2,039
Is self employed	0.109	0.122	0.394	2,039
Own savings	0.093	0.084	0.522	2,026
Savings (USD)	0.767	0.622	0.315	2,008
VSLA member	0.120	0.115	0.837	2,039
Risk tolerance	2.430	2.520	0.338	2,039
Partnered	0.674	0.673	0.959	2,039
Pro-women attitudes index	0.058	0.144	0.084	1,996
Decisions index	-0.023	0.059	0.150	2,039
Locus of control	2.017	1.995	0.656	2,020
Depression index (PHQ-9)	6.610	7.265	0.033	2,006
Anxiety index (GAD-7)	6.838	7.149	0.338	2,039
Physical health index (ADLs)	0.015	0.008	0.909	1,984

Notes. Consumption, earnings, and savings all winsorized at the 5th and 95th percentiles. Business costs include the discounted use value of large purchases. "Is Self Employed" = 1 if the respondent's primary job is self employment. "Partnered" = 1 if the respondent is living with a partner, whether formally married or not. Decisions, locus of control, pro-women attitudes and physical health indices constructed following Anderson (2008) and standardized so that the control group at endline has mean of zero and standard deviation of one.

6 Results

Table 2 shows the estimates of four meta-indices, one for each family of outcomes.[16] The average ITT effect sizes at the end of the program are 0.12SD for consumption, 0.08SD for employment and finances, 0.18SD for women's empowerment, and 0.04SD for health. These effects are generally as large in the second year, showing that the improvements outlive the intervention and do not fade out within this time interval.

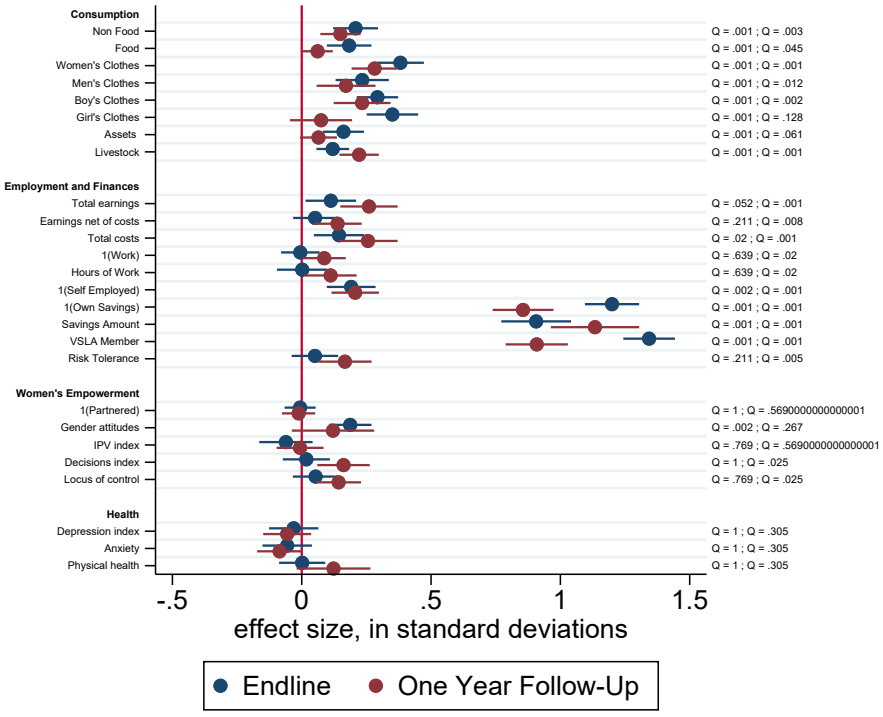
Table 2: Effects of the Pooled Treatments on Meta-Indices (ITT)

Dependent Variable	Index of			
	Consumption	Employment & Finances	Women Empowerment	Health
Panel A: Endline				
Treatment	0.125***	0.0859***	0.178***	0.0370
	(0.0293)	(0.0229)	(0.0597)	(0.0594)
	[.001]	[.001]	[.003]	[.155]
Control mean of dep. var	0	0	0	0
N	1,886	1,852	1,278	1,807
Panel B: 1 Year Follow-Up				
Treatment	0.100***	0.0804***	0.207***	0.0839
	(0.0327)	(0.0234)	(0.0631)	(0.0504)
	[.004]	[.004]	[.004]	[.026]
P-value: $\beta^{End} = \beta^{1YFU}$	0.896	0.372	0.776	0.697
Control mean of dep. var	0.0980	0	-0.0660	0.313
N	1,793	1,759	1,295	1,607

Notes. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Each index constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

[16]Specifically, we construct an index following Anderson (2008) for consumption, employment and finances, pro-women attitudes, and health by combining the outcomes in tables 3, 4, 5, and 6, respectively. The only exception is that because we only asked the IPV questions to partnered women, we defined the gender meta-index conditional on being currently partnered rather than including partnership status as an outcome in the index. Note that, as table 5 indicates, there are no treatment effects on remaining partnered, so there does not seem to be differential selection into having a report of the gender meta-index.

Figure 1: Treatment Effects at-a-glance



Notes: Sharpened q-values to the right of the estimates (Benjamini et al., 2006) control the false discovery rate; the first q-value pertains to the endline estimate and the second pertains to the one year follow-up. Asset, livestock, pro-women attitudes, IPV, decisions, and physical health indices all constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Figure 1 shows the treatment effects for each outcome, with the effect sizes in standard deviations of the control group at endline. We also show each non-standardized outcome in Tables 3 to 6.

Table 3 shows the estimated average effects on consumption and assets. We note that food consumption is approximately two-thirds of total budget, a hallmark of poverty and of similar magnitude as the food budget share for rural recipients of PROGRESA, Mexico's conditional cash transfer program (Angelucci and De Giorgi, 2009).

We find large increases in consumption of both non-durable and durable goods. Non-food expenditures increase by 20 and 13 percent at endline and one year follow-up. Food consumption

increases by 15 percent at endline and 5 percent at follow-up. The asset and livestock indices increase by 0.16 and 0.12SD at endline and 0.06SD and 0.22SD at follow-up, respectively.

These findings have two important implications. First, to the extent that consumption is an indicator of poverty, these large, positive, and persistent impacts suggest that the intervention has substantially reduced household poverty and that the program continues to reduce poverty two years after its start. Second, these large increases in consumption indicate that the program generates positive spillover effects to the entire household. Therefore, considering only how the program benefits its direct recipients would largely underestimate the overall impacts of this intervention (Angelucci and De Giorgi, 2009). Consistent with the idea of positive spillover effects in the households, we find that school enrollment for children aged 5 to 18 increases by 5 percentage points from a control mean of 71 percent (p-value 0.006).[17]

We also find large and positive impacts on clothing purchases, expenditures on which almost double at endline; we find sustained positive impacts, though smaller, at follow-up. Since we have assignable goods, we can compute the budget shares for women's and men's clothing. We find that, at endline, the program increases the budget share of women's clothing by 56 percent, from 2.5 to 3.9 percent and the budget share of men's clothing by 50 percent, from 0.4 to 0.6 percent.[18] At follow-up, the program increases the budget share of women's clothing by 38 percent, from 2.6 to 3.6 percent and the budget share of men's clothing by 20 percent, from 0.5 to 0.6 percent.[19] This suggests that the program may have increased women's bargaining power, especially at follow-up. We revisit this issue below in our discussion of women's empowerment.

Table 4 shows impacts on employment and financial outcomes. Unless mentioned otherwise, we focus on the treatment effects at follow-up. We find that the program has statistically and economically significant impacts in this domain. Specifically, the program increases both earnings (gross and net) and business costs. The higher business costs suggest increases or expansion of entrepreneurial activities. Net earnings increase by about 20 percent, suggesting that there is also an increase in profit. Similarly, self-employment increases from 12 to 19 percent. These increases in entrepreneurship are partly due to increased entry into employment and switches from wage- to self-employment: the probability of working for pay or profit increases by 4.3 percentage points in the treatment group at follow-up, a 10 percent increase. Similarly, unreported tabulations show

[17]Results available upon request. We omit this outcome from the result tables because we did not pre-specify it in our pre-analysis plan.

[18]For example, the budget shares of women's clothing is $\frac{0.316}{4.076+8.431}$ =2.5 percent in the control group at endline.

[19]Unreported regressions confirm that these changes are statistically significant.

that the probability of being self-employed outside of agriculture increases from 18 to 30 percent, conditional on having a job. Hours of work in the previous week similarly increased by 1.55 hours, a 19 percent increase.[20]

We also find that the program has a large and positive impact on women's savings. The fraction of women with monetary savings at follow-up doubles, increasing by 37 percentage points from 36 percent of the control group. Savings balances almost triple, significantly increasing from USD 4.8 to 13.0. A channel for higher savings is VSLA membership, which also doubles from 38 to almost 80 percent. Lastly, women's attitudes to risk also change: we find that the program increases risk tolerance by about 10 percent. This finding is consistent with, and may partly explain, the higher rates of entrepreneurship.

Comparing the magnitude of the treatment effects at endline and follow-up shows that these impacts are not fading within the evaluation time frame, and several of them, such as gross earnings and personal savings, are actually increasing. These effects suggest that our findings are not driven only by the cash transfer. Other program features, such as training and empowerment (which we discuss next), may also increase women's earnings capacity and financial resources.

A key consideration in interpreting the effects on women's income is what happens to the income of other household members. We do this by examining treatment effects on partners' income at endline.[21] Note that we have information only on partners who are primary residents of the household (the qualification to appear in the roster, where earnings questions were asked); this is true of 85 percent of partners of women who were partnered at endline. Table A3 assesses selection into spouse presence among women partnered at endline and shows that there is no differential effect by treatment. Table A4 then estimates treatment effects on partner earnings in this sample. Similar to Bernhardt et al. (2019), we find that husbands' earnings increase by USD 0.72 per week, or 62 percent. This increase is statistically higher than the impact on wives' earnings ($P = 0.060$), which increase by USD 0.2, or 19 percent.

[20]Table A5 estimates the treatment effect on time use in the previous 24 hours. It is likely that time use over the past 24 hours is noisier than labor supply measured over the past seven days. Indeed, we find no significant effect on reported minutes in market labor over the past day.

[21]A coding issue in the roster for the one year follow-up led to missing values for spousal characteristics, including income. Thus, we do not attempt to estimate treatment effects on partners' income at follow-up.

Table 3: Effects of the Pooled Treatments on Weekly Non-Durable Consumption, Assets, and Livestock (ITT)

Dependent Variable	Total Consumption		Clothing Expenditure on				Durables (Indices)	
	Non-Food	Food	Women's	Men's	Girls'	Boys'	Assets	Livestock
Panel A: Endline								
Treatment	0.804***	1.086***	0.210***	0.0377***	0.0443***	0.0375***	0.158***	0.122***
	(0.201)	(0.307)	(0.0300)	(0.00999)	(0.00728)	(0.00641)	(0.0464)	(0.0386)
	[.052]	[.211]	[.02]	[.639]	[.639]	[.002]	[.001]	[.001]
Control mean of dep. var	4.076	8.431	0.316	0.0480	0.0590	0.0390	-0.00600	0.00600
N	1,887	1,887	1,887	1,887	1,887	1,887	1,886	1,887
Panel B: 1 Year Follow-Up								
Treatment	0.577***	0.364*	0.155***	0.0276**	0.0353***	0.00800	0.0634	0.225***
	(0.180)	(0.209)	(0.0294)	(0.0111)	(0.00999)	(0.00771)	(0.0418)	(0.0463)
	[.001]	[.008]	[.001]	[.02]	[.02]	[.001]	[.001]	[.001]
P-value: $\beta^{End} = \beta^{1YFU}$	0.087	0.204	0.226	0.191	0.202	0.964	0.000	0.031
Control mean of dep. var	4.458	7.397	0.305	0.0590	0.0880	0.0670	0.00800	0.00700
N	1,793	1,793	1,793	1,793	1,793	1,793	1,793	1,793

Notes. All consumption values in USD; durables indices constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table 4: Effects of the Pooled Treatments on Employment and Finances (ITT)

Dependent Variable	Total earning (USD)	Earnings net of costs (USD)	Total business costs (USD)	Worked last week	Hours of work last week	Self employed	Own savings	Savings (USD)	VSLA member	Risk tolerance
Panel A: Endline										
Treatment	0.202*	0.0714	0.180**	-0.00269	0.0266	0.0626***	0.519***	6.567***	0.602***	0.0884
	(0.106)	(0.0704)	(0.0731)	(0.0223)	(0.814)	(0.0185)	(0.0273)	(0.588)	(0.0269)	(0.0936)
	[.052]	[.211]	[.02]	[.639]	[.639]	[.002]	[.001]	[.001]	[.001]	[.211]
Control mean of dep. var	1.081	0.804	0.381	0.446	8.187	0.122	0.249	2.920	0.278	2.627
N	1,887	1,879	1,879	1,887	1,887	1,887	1,869	1,835	1,887	1,887
Panel B: 1 Year Follow-Up										
Treatment	0.467***	0.191**	0.321***	0.0433*	1.548*	0.0677***	0.370***	8.215***	0.407***	0.289***
	(0.120)	(0.0773)	(0.0859)	(0.0249)	(0.837)	(0.0180)	(0.0306)	(0.744)	(0.0324)	(0.107)
	[.001]	[.008]	[.001]	[.02]	[.02]	[.001]	[.001]	[.001]	[.001]	[.005]
P-value: $\beta^{End} = \beta^{1YFU}$	0.087	0.204	0.226	0.191	0.202	0.964	0.000	0.031	0.000	0.133
Control mean of dep. var	1.319	0.963	0.378	0.463	8.235	0.122	0.355	4.832	0.381	2.446
N	1,793	1,786	1,786	1,793	1,793	1,793	1,779	1,755	1,793	1,793

Notes. Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

This finding is important for multiple reasons. First, it shows that there are positive spillover effects of the treatment on spousal earnings. Focusing on the treatment effects on the treated women, therefore, substantially underestimates the overall program impact. Second, there may be multiple pathways through which the intervention increases spouses' incomes. With fixed costs in production, economies of scale in the household may generate positive spillovers to partners' income. Alternatively, spouses' enterprises may be more profitable (Bernhardt et al., 2019) and, therefore, investing in them may be efficient. In addition, some spouses may appropriate their wives' resources. These pathways may impact women's wellbeing differently. Changes in relative incomes may also reflect changes in wives' and husbands' bargaining power. If the program increases both spouses' income, it may increase husbands' and wives' bargaining power relative to other household members and each other. We return to this issue in section 8.2.

Table 5: Effects of the Pooled Treatments on Women Empowerment (ITT)

Dependent Variable	Partnered	Pro-women attitudes index	IPV index	Decisions index	Locus of control
Panel A: Endline					
Treatment	-0.00286 (0.0164) [1]	0.187*** (0.0494) [.002]	-0.0976 (0.0987) [.769]	0.0179 (0.0547) [1]	0.0538 (0.0529) [.769]
Control mean of dep. var N	0.714 1,887	0 1,828	0.0450 1,326	0 1,887	2.006 1,871
Panel B: 1 Year Follow-Up					
Treatment	-0.00540 (0.0174) [.569]	0.121 (0.0957) [.267]	-0.0102 (0.0867) [.569]	0.162*** (0.0609) [.025]	0.143*** (0.0520) [.025]
P-value: $\beta^{End} = \beta^{1YFU}$	0.640	0.476	0.497	0.053	0.196
Control mean of dep. var N	0.755 1,793	-0.0490 1,751	0.00600 1,329	-0.0940 1,793	2.017 1,773

Notes. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Pro-women attitudes, decisions, and IPV indices constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one.Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes nonmissing), PHQ-9 depression index, 1(PHQ-9 depression index nonmissing). Standard errors clustered at the level of treatment (group).

Table 5 estimates the treatment effects on Women's empowerment. We generally find improve-

ments on all dimensions considered, although not all estimated effects are individually statistically significant. Pro-women attitudes, women's participation in household decision-making, and locus of control scores increase, with most impacts still significant at follow-up.[22] The gains in household decision-making, as shown in Table A6, seem largely driven by women's increased participation in decisions about income generation, which aligns with findings in Bossuroy et al. (2021). IPV decreases, but this effect is small and statistically insignificant. These impacts suggest that the treatment increases women's empowerment by increasing their autonomy, their sense of control over their lives, and their ideas of the role of women in society. These increases in empowerment are not inconsistent with the result that partners' earnings increase more than treated women's earnings: the program still plausibly increased both women's outside options and their control of income (as evidenced by increased rates of own savings shown in table 4).

The lack of significant reductions in average IPV is not inconsistent with these findings. IPV largely depends on partners' beliefs and behavior. Moreover, the program may have heterogeneous effects on IPV, causing it to increase among some couples and decrease in others. We return to this issue in Section 8.2.

Table 6 shows the estimated treatment effects on mental and physical health. By follow-up, the treatment effect on each measure of health improves, though none is individually statistically significant. The lack of large or statistically significant health improvements despite large increases in consumption and expenditures suggests that other features of the program may have offset the benefits of higher and better nutrition.[23] For example, women's health may not improve if the program increases physical exertion and stress, reduces leisure time, or increases IPV for some participants.[24] This is one important group of outcomes for which this program might have unintended effects. These findings contrast with the impacts of graduation programs that do not target women (Banerjee et al., 2015), which find modest positive effects on participants' physical and mental health (0.032 SD and 0.099 SD), and with the interventions that target women, which find positive effects that range from 0.077 SD to 0.23 SD.[25]

[22]Unreported regressions also find statistically significant increases in women's social connectedness, as measured by current participation in a community group.

[23]Unreported regressions find statistically significant improvements in household diet diversity scores, to be explored in a future paper.

[24]The lack of health care infrastructure or other supply-side constraints may also contribute to this muted effect.

[25]From Bandiera et al. (2017) and Bossuroy et al. (2021). Bedoya et al. (2019) do not measure participants' health.

Table 6: Effects of the Pooled Treatments on Mental and Physical Health (ITT)

Dependent Variable	Depression index (PHQ-9)	Anxiety index (GAD-7)	Physical health index (ADLs)
Panel A: Endline			
Treatment	-0.144 (0.266) [1]	-0.283 (0.288) [1]	0.00170 (0.0535) [1]
Control mean of dep. var	7.174	6.942	0
N	1,846	1,887	1,836
Panel B: 1 Year Follow-Up			
Treatment	-0.262 (0.259) [.305]	-0.430 (0.262) [.305]	0.123 (0.0858) [.305]
P-value: $\beta^{End} = \beta^{1YFU}$	0.845	0.810	0.311
Control mean of dep. var	6.081	6.750	-0.0130
N	1,753	1,793	1,699

Notes. ADLs = Activities of Daily Living. ADL index constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one.Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes nonmissing), PHQ-9 depression index, 1(PHQ-9 depression index nonmissing). Standard errors clustered at the level of treatment (group).

7 Cost-Benefit Analysis

The total direct and indirect costs of the program's full implementation are USD 354 per participant.[26] We compare these costs with the program impacts on non-durable consumption at endline and one year follow-up. These are USD 109 at endline and USD 63 at follow-up.[27] We estimate the present value of the effects of the program on nondurable consumption. This choice is conservative

[26]This cost includes all in-country costs for program delivery including indirect costs and excludes technical assistance from headquarters staff.

[27]This is the weekly value from table 3 divided by 7 and multiplied by 365 to obtain the annual value. We then take this value, which is in 2017 USD, and convert it into 2021 USD by multiplying by 1.1

because it excludes the program impacts on savings, assets, the other outcomes we considered, and other spillover effects in the household.[28]

Using an annual discount rate of 0.05 and assuming that the one-year follow-up benefits last in perpetuity, as in Banerjee et al. (2015), the present value of the intervention is USD 1306. The intervention breaks between 4 and 5 years. The benefit-cost ratio (BCR) is 368 percent, and the Internal Rate of Return (IRR) is 19.9 percent. These ratios are towards the upper end of the estimates in Banerjee et al. (2015). By these estimates, the program is successful.[29]

8 Heterogeneity

8.1 Heterogeneity by baseline disadvantage

Following our pre-analysis plan, we consider subgroup-specific effects along ten dimensions, as captured at baseline: literacy, employment status, whether the respondent was partnered, depression, anxiety, decision-making index, pro-women attitudes, predicted IPV, spousal age gap, and the respondent being the primary household earner.

The findings, shown in Table A7, hint at more disadvantaged women benefiting more from the intervention. For example, the increase in market labor at the one year follow-up is higher for women with higher baseline depression, with a larger spousal age gap, and with lower decision-making power. Similarly, IPV decreases more for illiterate women and women with a higher spousal age gap, and physical health increases more for women with lower decision-making power, at higher IPV risk, and who were not the primary household earner. While one hopes is that more disadvantaged women will benefit disproportionately from anti-poverty programs, it is *ex ante* unclear that they will, as the support package offered may be not big enough to overcome the barriers to personal and economic development that these women face. However, having 26 outcomes, two time periods, and ten interactions for each makes it difficult to summarize succinctly.

To offer findings on heterogeneity more concisely, we estimate treatment effects by baseline socioeconomic disadvantage. To account for the multidimensional nature of socioeconomic disadvantage, we consider four indices: IPV risk, economic status, health, and social connectedness. Each index is standardized such that higher values indicate higher SES. Table A8 provides further

[28]Since the program increases partners' income, it is likely that the effect on household savings is greater than the effect on participants' savings, since it is unlikely that all additional spousal income is consumed.

[29]We consider two alternative scenarios in which the one-year follow-up benefits last for 20 or 10 years. In these cases, the present value is USD 845 and USD 516, and the IRR is 19.4 percent and 15.2 percent, respectively.

details. The correlation coefficient between them ranges from 0.006 to 0.187, suggesting that these four indices capture distinct aspects of disadvantage. We then interact the treatment with a dummy variable indicating an above-median value on each index.

To test whether the treatment effects vary systematically along these four dimensions of disadvantage, we estimate all treatment effects as a system of equations and test whether the interaction terms are jointly statistically significant across all outcomes and indices. Table A8 provides the detailed estimates for each outcome and index. We reject the null hypotheses that the interactions are jointly zero ($P < 0.001$), suggesting that the treatment effects vary systematically with disadvantage. However, we find no differential effects on household-level outcomes such as consumption and assets or any clear pattern for the individual-level outcomes (e.g., that the program benefits disadvantaged women more). One notable exception is the effect on IPV, which increases for women at low baseline-IPV risk and decreases for women at high baseline-IPV risk. We investigate this in the next section.

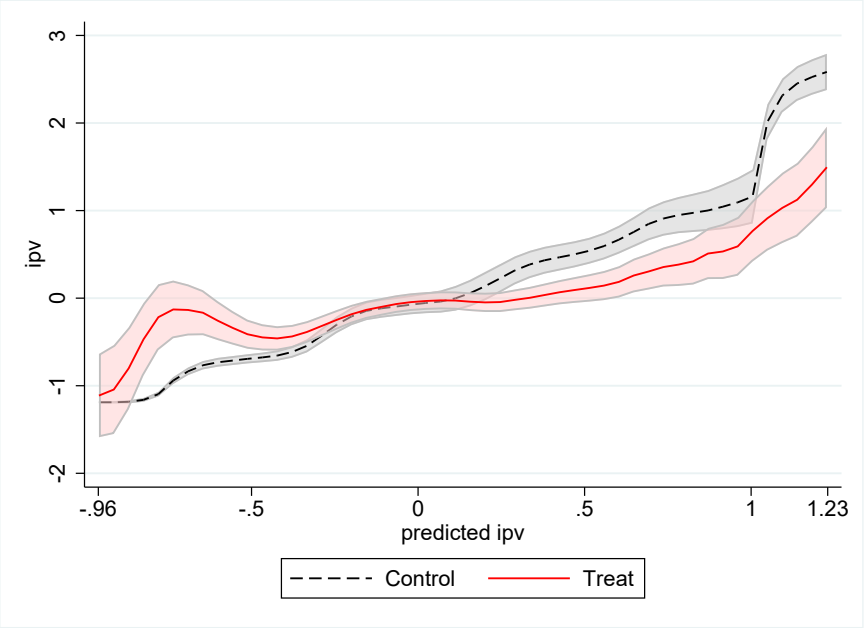
8.2 Empowerment and IPV

The relationship between empowerment and IPV is complex. If IPV is a normal good for partners, women's empowerment can increase their bargaining power, thus reducing equilibrium IPV. At the same time, women's higher wealth and income may increase the returns of instrumental IPV (i.e., partners' acts of violence meant to gain control of spousal resources). Lastly, if partners' identities feel threatened by the heightened women's status, the men may retaliate by increasing IPV to assert their dominance. Consequently, we expect treatment effects to vary acrosss couples. This is consistent with our findings in section 8.1.

To further investigate the heterogeneous treatment effects on IPV, we create a standardized IPV risk index using a Random Forest model and a large set of predetermined explanatory variables. Appendix A.1.4 provides further details.

Figure 2 shows how the severity of IPV at endline and follow-up (on the y-axis) varies non-parametrically with baseline IPV risk (on the x-axis) for women in the treatment and control groups, pooling across treatment arms and rounds. Higher values of the index correspond to higher baseline IPV risk. Consistent with our previous results, we find that the intervention reduces IPV for women with IPV risk of 0.2SD or higher. This group is approximately 50 percent

Figure 2: IPV prevalence by treatment status and IPV risk predicted by baseline characteristics – Pooled across treatments and rounds



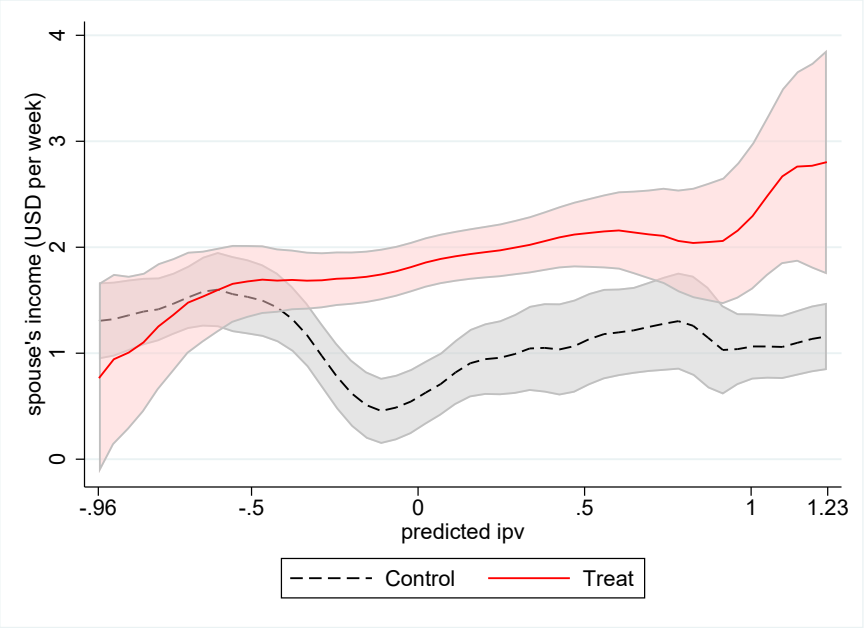
Notes: IPV (vertical axis) is constructed as described in section 3.3 and standardized so that the control group at endline has mean zero and standard deviation one. Predicted IPV (horizontal axis) is constructed as described in section 8.2.

of the sample. Conversely, the treatment increases IPV severity for women with IPV risk of -0.3SD and lower. This group is approximately 25 percent of the sample.

Since one of the motives for IPV is to seize control of the woman's resources, we investigate if the effects on IPV and partners' incomes are correlated. Figure 3 indicates that the increase in husbands' income is concentrated among women with predicted IPV of -0.2 and higher. This is the group who also experiences no change or a reduction in IPV. Conversely, among women with a predicted IPV of -0.5 or less, we find both increases in IPV and no change in husbands' income. We conjecture that the two effects may be related, either because the program fosters cooperation among the former group of households, or because the women use part of their stipend to reduce IPV, likely transferring some resources to their partners.

While heterogeneity based on factors such as partner characteristics (Angelucci, 2008) or family structure (Heath et al., 2020) is commonly found in the relationship between cash transfers and IPV, we are not aware of other interventions that have been found to reduce IPV among high-risk

Figure 3: Partner Income by treatment status and IPV risk



Notes: Predicted IPV (horizontal axis) is constructed as described in section 8.2

women.[30]

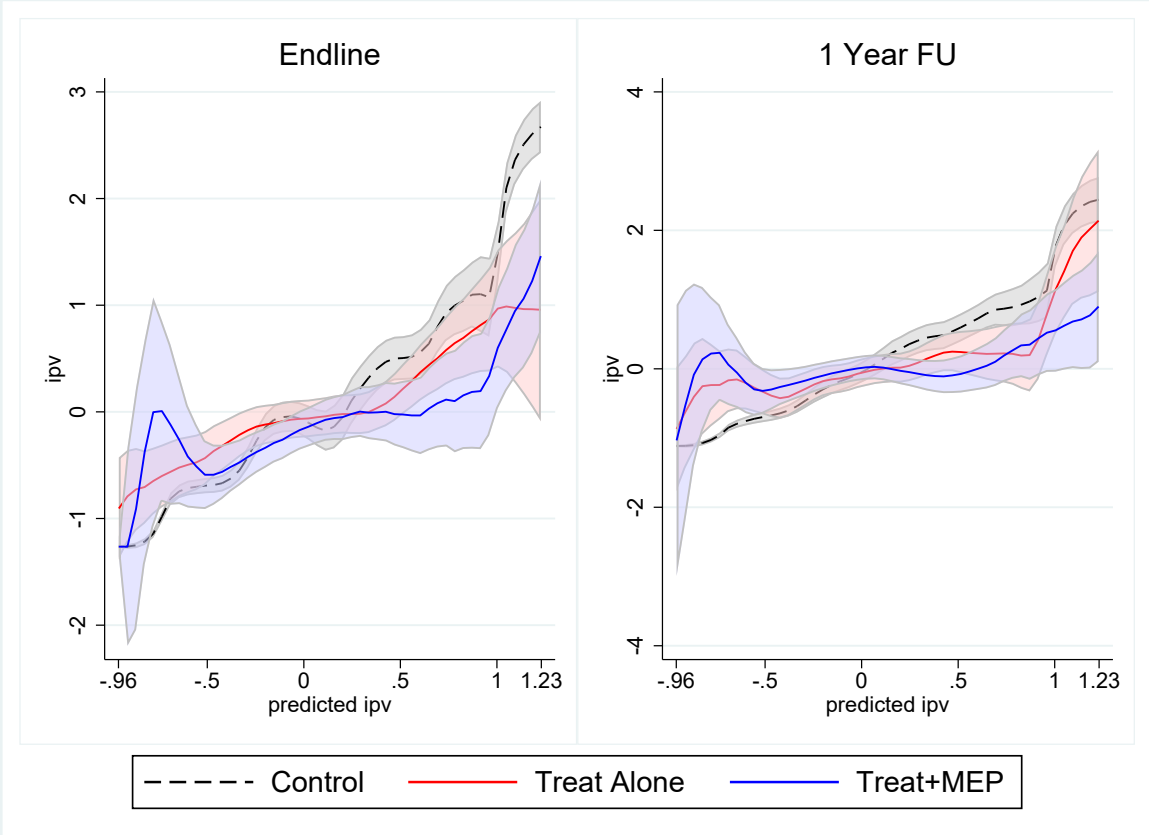
Figure 4 shows the relationship between IPV intensity and IPV risk at baseline separately by arm (women-only intervention vs women's intervention + MEP) and period. There are two main findings. First, the differences from the control group are more marked for the MEP group. However, the levels of IPV in general do not differ statistically between these two arms. Second, the small IPV increase for women at lower IPV risk is more marked at the one year follow-up. This finding suggests that the increase of IPV may not be a temporary phenomenon and, in fact, could outlasts the intervention.

9 The effects of the MEP program

Figure 5 show the estimated differential treatment effects for the MEP intervention, estimated by adding an interaction term for treatment × MEP in equation 1. At endline, the MEP treatment has negative differential effects on IPV and positive differential effects on gender attitudes. However, neither coefficient is statistically significant once we correct for multiple inference, and the effects

[30]See Buller et al. (2018) for a review of studies on the relationship between cash transfers and IPV, which includes a discussion of common dimensions of heterogeneity.

Figure 4: IPV prevalence by treatment status and IPV risk predicted by baseline characteristics – By treatment and round



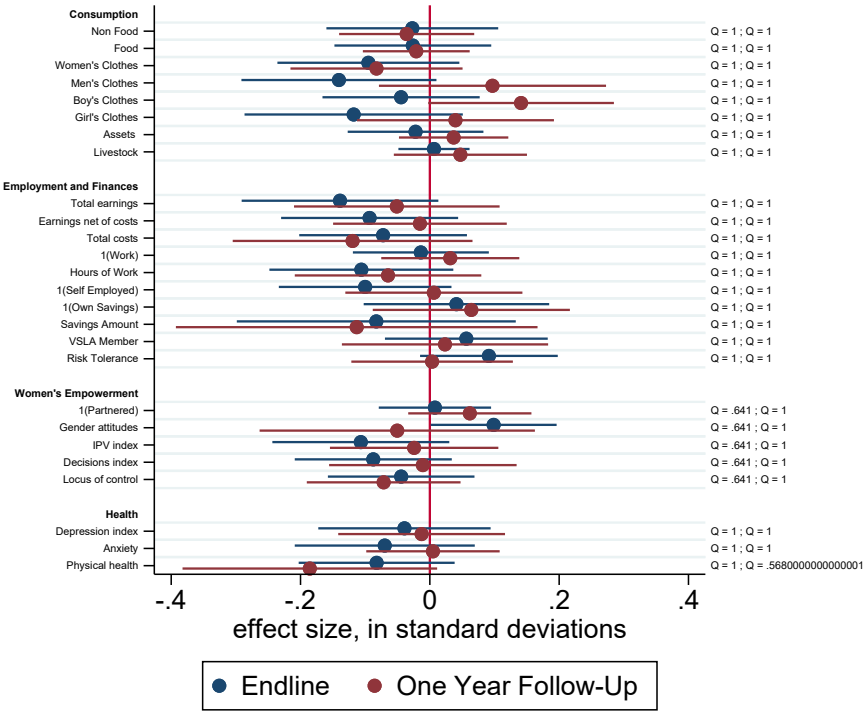
Notes: IPV (vertical axis) is constructed as described in section 3.3 and standardized so that the control group at endline has mean zero and standard deviation one. Predicted IPV (horizontal axis) is constructed as described in section 8.2

fade over time. These and other relevant study findings (e.g., Vaillant et al. (2020) suggest that light-touch interventions for men such as the one implemented in this trial may not be sufficient to facilitate additional gains in women's outcomes beyond those due to the women's intervention in the longer term.

10 Discussions and Concluding Remarks

This study evaluated a graduation program that explicitly targets women and seeks to increase their socioeconomic power. The paper shows that it is possible to empower ultra-poor women in war-torn settings with high gender inequality. Since the intervention's IRR is as high as or higher than those of other graduation programs that do not directly target women, these results

Figure 5: ITT estimates of the MEP program



Notes: Sharpened q-values to the right of the estimates (Benjamini et al., 2006) control the false discovery rate; the first q-value pertains to the endline estimate and the second pertains to the one year follow-up. Asset, livestock, pro-women attitudes, IPV, decisions, and physical health indices all constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

further suggest that empowering women within graduation interventions can be achieved in a cost-effective way.

We note that the program leads to positive spillover effects in the household on both durable and non-durable non-food expenditures and food consumption. These positive spillovers in the household are a common feature of anti-poverty programs targeting women, and they are one of the reasons why women are the recipients of cash transfer programs in many countries (Armand et al., 2016; Yoong et al., 2012). In addition, we find positive spillovers on children's schooling, which increases by 5 percentage points, and partners' income, which increases four times as much as women's income. We conclude that our estimates of the program impacts on income and finances,

which focus on the effects on the recipients, may underestimate the total effect of the program.

We discuss how programs that successfully improve women's livelihoods and status may face additional constraints, depending on societal norms, and we explore possible backlash. Consistent with theory, we find heterogeneous effects on IPV: while there is a small and statistically insignificant reduction in average rates of IPV, the program reduces self-reported IPV for women at higher-than-average risk for violence, but it may increase IPV for women at lower risk.

While attention is needed to help counteract potential program-induced backlash against some women, the positive effects of the intervention on women in a very poor, post-conflict setting is a positive sign for policymakers interested in improving women's welfare.

References

Adoho, Franck M and Djeneba Doumbia, *Informal sector heterogeneity and income inequality: Evidence from the Democratic Republic of Congo*, The World Bank, 2018.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Angelucci, Manuela, "Love on the rocks: Domestic violence and alcohol abuse in rural Mexico," *The BE Journal of Economic Analysis & Policy*, 2008, *8* (1).

_ **and Giacomo De Giorgi**, "Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption?," *American Economic Review*, 2009, *99* (1), 486–508.

Armand, Alex, Orazio Attanasio, Pedro Manuel Carneiro, and Valérie Lechene, "The effect of gender-targeted conditional cash transfers on household expenditures: Evidence from a randomized experiment," *CEPR Discussion Paper No. DP11465*, 2016.

Attanasio, Orazio and Valerie Lechene, "Tests of income pooling in household decisions," *Review of Economic Dynamics*, 2002, *5* (4), 720–748.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, "Labor markets and poverty in village economies," *The Quarterly Journal of Economics*, 2017, *132* (2), 811–870.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry, "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, 2015, *348* (6236).

Bedoya, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro, "No household left behind: Afghanistan targeting the ultra poor impact evaluation," Technical Report, National Bureau of Economic Research 2019.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, 2006, *93* (3), 491–507.

Bernhardt, Arielle, Erica Field, Rohini Pande, and Natalia Rigol, "Household matters: Revisiting the returns to capital among female microentrepreneurs," *American Economic Review: Insights*, 2019, *1* (2), 141–60.

Bjørkhaug, Ingunn and Morten Bøås, *Men, women, and gender-based violence in North Kivu, DRC*, Fafo, 2014.

Bossuroy, Thomas, Markus Goldstein, Dean Karlan, Harounan Kazianga, William Pariente, Patrick Premand, Catherine Thomas, Christopher Udry, Julia Vaillant, and Kelsey Wright, "Pathways out of Extreme Poverty: Tackling Psychosocial and Capital Constraints with a Multifaceted Social Protection Program in Niger," *World Bank, Washington, DC*, 2021.

Buller, Ana Maria, Amber Peterman, Meghna Ranganathan, Alexandra Bleile, Melissa Hidrobo, and Lori Heise, "A mixed-method review of cash transfers and intimate partner violence in low- and middle-income countries," *The World Bank Research Observer*, 2018, *33* (2), 218–258.

Coghlan, Benjamin, Pascal Ngoy, Flavien Mulumba, Colleen Hardy, Valerie Nkamgang Bemo, Tony Stewart, Jennifer Lewis, and Richard J Brennan, "Update on mortality in the Democratic Republic of Congo: results from a third nationwide survey," *Disaster Medicine and Public Health Preparedness*, 2009, *3* (2), 88–96.

Conceição, Pedro, *Human Development Report 2019: beyond income, beyond averages, beyond today: inequalities in human development in the 21st century*, United Nations Development Programme, 2019.

Davis, Laura, Paola Fabbri, and Ilot Muthaka Alphonse, "Democratic Republic of Congo–DRC: gender country profile 2014," 2014.

Gibbs, Andrew, Julienne Corboz, Esnat Chirwa, Carron Mann, Fazal Karim, Mohammed Shafiq, Anna Mecagni, Charlotte Maxwell-Jones, Eva Noble, and Rachel Jewkes, "The impacts of combined social and economic empowerment training on intimate partner violence, depression, gender norms and livelihoods among women: an individually randomised controlled trial and qualitative study in Afghanistan," *BMJ Global Health*, 2020, *5* (3), e001946.

Glinski, A, C Schwenke, L O'Brien-Milne, and K Farley, "Gender equity and male engagement: It only works when everyone plays," *Washington, DC, ICRW*, 2018.

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, 2012, *20*, 25–45.

_ **and Yiqing Xu**, "ebalance: A Stata Package for Entropy Balancing," *Journal of Statistical Software*, 2013, *54* (7).

Heath, Rachel, Melissa Hidrobo, and Shalini Roy, "Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali," *Journal of Development Economics*, 2020, *143*, 102410.

Kroenke, Kurt, Robert L Spitzer, and Janet BW Williams, "The PHQ-9: validity of a brief depression severity measure," *Journal of General Internal Medicine*, 2001, *16* (9), 606–613.

MPSMRM, Ministère de la Santé Publique (MSP), and ICF International, "Enquête démographique et de santé en République Démocratique du Congo 2013–2014," 2014.

Mughal, Anisa Y, Jackson Devadas, Eric Ardman, Brooke Levis, Vivian F Go, and Bradley N Gaynes, "A systematic review of validated screening tools for anxiety disorders and PTSD in low to middle income countries," *BMC Psychiatry*, 2020, *20* (1), 1–18.

Rotter, Julian B, "Generalized expectancies for internal versus external control of reinforcement.," *Psychological Monographs: General and Applied*, 1966, *80* (1), 1.

Spitzer, Robert L, Kurt Kroenke, Janet BW Williams, and Bernd Löwe, "A brief measure for assessing generalized anxiety disorder: the GAD-7," *Archives of Internal Medicine*, 2006, *166* (10), 1092–1097.

Vaillant, Julia, Estelle Koussoubé, Danielle Roth, Rachael Pierotti, Mazeda Hossain, and Kathryn L Falb, "Engaging men to transform inequitable gender attitudes and prevent intimate partner violence: a cluster randomised controlled trial in North and South Kivu, Democratic Republic of Congo," *BMJ Global Health*, 2020, *5* (5), e002223.

World Health Organization et al., "Understanding and addressing violence against women: Intimate partner violence," Technical Report, World Health Organization, 2012.

Yoong, Joanne, Lila Rabinovich, and Stephanie Diepeveen, "The impact of economic resource transfers to women versus men: a systematic review," *Institute of Education technical report, University of London (London, EPPI-Centre)*, 2012.

Zeender, Greta and Jacob Rothing, "Displacement trends in DRC," *Forced Migration Review*, 2010, (36), 10.

A Appendix (For Online Publication Only)

A.1 Disadvantage Indices

We construct four indices of socioeconomic disadvantage at baseline. Table A9 provides the correlation between each index.

A.1.1 Economic Disadvantage

We construct an index following Anderson (2008), using the following variables, all defined at baseline:

- Respondent's education, amount of savings, a dummy for any savings, and net earnings
- Partner's education, earnings, and age
- Presence of shocks to the household in the previous twelve months: illness lasting at least one month, death of a household member, a period of at least one month when someone couldn't find work, loss of a business, loss of a significant part of household assets (examples of losses were provided: animals were stolen or died, goods were stolen or burned or are noo longer usable), significant increase in prices of daily goods, a divorce or separation, displacement of some or all household members, incident of violence (community level or individual), some other serious loss

For any variables with some missing values, we replace missing values with zeros and include a dummy variable for whether the variable is missing.

A.1.2 Physical and Mental Health

We construct an index following Anderson (2008), using the following variables, all defined at baseline and inverted so that higher values indicate better health:

- depression
- anxiety
- physical health (ADLs)

A.1.3 Social Connectedness

We construct an index following Anderson (2008), using the following variables, all defined at baseline:

- Whether the respondent is currently a member, participant, volunteer or otherwise interacts with any of the following types of groups (binary)
  - a dance, music, drama or other cultural group,
  - a cooperative, farmers group or business group,
  - a women's group,
  - a political party or political group,
  - a church or mosque group or any other religious group, or

  - any other type of group not mentioned so far

- Whether the respondent has someone who can give her money when she needs it (binary)
- Whether the respondent has a place that she could go to sleep if there is an emergency or something happened to make her feel unsafe where she usually sleeps
- How many times in the past four weeks she has gone to someone else to discuss a problem she is facing or something else important
- How many times in the past four weeks someone has come to her to discuss a problem they are facing or something else important
- How many people she considers friends that she spent time with in the last week, not including household members

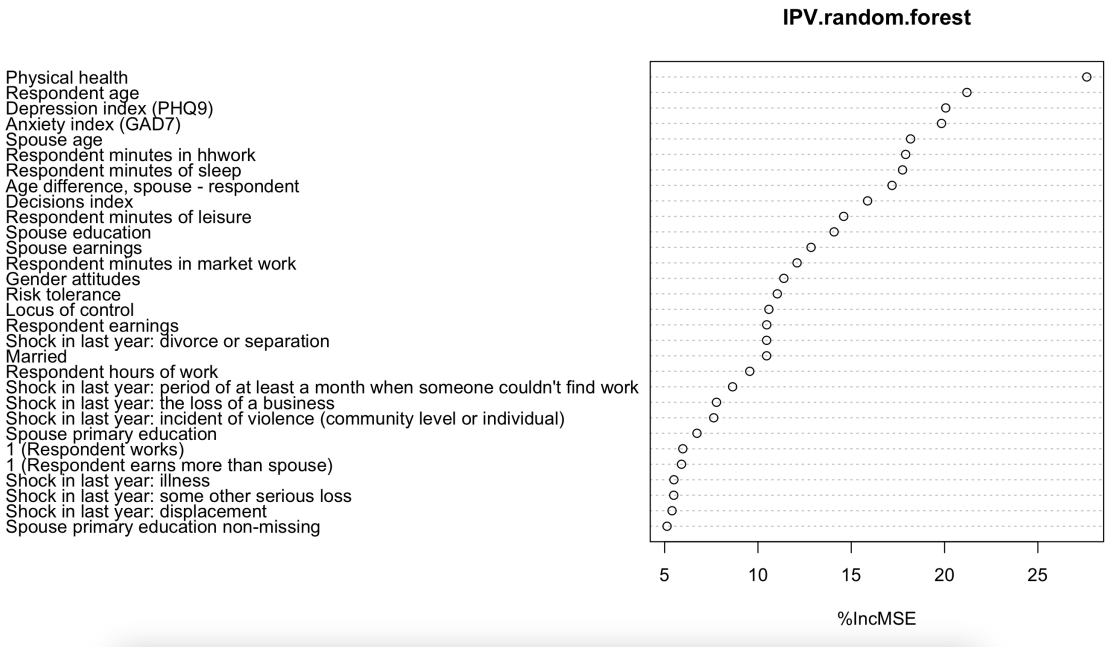
A.1.4 IPV Risk

We use predicted IPV based on a Random Forest model using only the control group (at endline and one year follow-up), and then use estimated coefficients from the model to generate predictions in the treated group as well. The covariates we use include baseline values of our outcomes (in tables 3, 4, 5, 6, and A5), as well as variables that theory highlights as risk factors for IPV: those that capture recent stress (economic shocks), the outside options of both spouses (each spouse's age, education, and earnings at baseline), and the potential for backlash from the male partner (a binary variable capturing whether the wife earns more).

Tree models recursively partition the data based on a decision rule where some covariate X is less than some constant. The constant c is selected such that the information gain from the partition is maximized. Here, we make use of a random forests which takes bootstrapped samples of our data and fits a tree to each one. The predictions across trees are then averaged. The random forest allows us to determine the most important variables for prediction where importance is defined by how much the mean squared prediction error (MSE) increases if a particular variable is excluded. Figure A1 plots variable importance by this metric, defined by how much the mean squared prediction error (MSE) increases if a particular variable is excluded.

A.2 Appendix Table and Figures

Figure A1: Variable importance in the random forest IPV prediction



Notes: 500 trees used in estimation. All variables are defined at baseline.

Table A1: estimates with entropy weights (ITT)

Dependent Variable	Endline			1YFU		
	β	SE	q-value	β	SE	q-value
Non-food consumption (USD)	0.829	0.205	0.001	0.452	0.206	0.056
Food consumption (USD)	1.102	0.312	0.001	0.361	0.207	0.086
Women's clothes (USD)	0.227	0.029	0.001	0.148	0.031	0.001
Men's clothes (USD)	0.040	0.010	0.001	0.026	0.012	0.056
Girls' clothes (USD)	0.048	0.007	0.001	0.035	0.010	0.003
Boys' clothes (USD)	0.041	0.006	0.001	0.009	0.008	0.169
Assets (standardized)	0.134	0.052	0.006	0.066	0.044	0.100
Livestock (standardized)	0.130	0.036	0.001	0.232	0.045	0.001
Total earnings (USD)	0.178	0.110	0.105	0.475	0.121	0.001
Earnings net of costs (USD)	0.067	0.073	0.224	0.185	0.080	0.014
Total business costs (USD)	0.160	0.079	0.061	0.331	0.083	0.001
Worked last week	-0.009	0.023	0.438	0.048	0.025	0.026
Hours of work last week	-0.014	0.836	0.652	1.729	0.843	0.020
Is self employed	0.064	0.019	0.002	0.071	0.018	0.001
Own savings	0.525	0.027	0.001	0.375	0.031	0.001
Savings (USD)	6.599	0.592	0.001	8.257	0.742	0.001
VSLA member	0.610	0.027	0.001	0.413	0.032	0.001
Risk tolerance	0.086	0.093	0.224	0.243	0.107	0.014
Partnered	-0.001	0.016	1.000	-0.007	0.017	0.373
Pro-women attitudes index	0.187	0.049	0.002	0.146	0.099	0.171
IPV index	-0.077	0.101	1.000	-0.051	0.094	0.373
Decisions index	0.012	0.054	1.000	0.167	0.062	0.023
Locus of control	0.045	0.055	1.000	0.145	0.053	0.023
Depression index (PHQ-9)	-0.227	0.273	1.000	-0.301	0.278	0.207
Anxiety index (GAD7)	-0.397	0.292	1.000	-0.427	0.268	0.207
Physical health index (ADLs)	0.010	0.055	1.000	0.149	0.087	0.207

Notes: Sharpened q-values (Benjamini et al., 2006) control the false discovery rate. Asset, livestock, pro-women attitudes, IPV, decisions, and physical health indices all constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A2: Attrition

Dependent Variable	Appear at Endline		Appear in the 1 Year Follow-Up	
Treatment	-0.00303	-0.000792	0.00462	0.00582
	(0.0157)	(0.0138)	(0.0197)	(0.0142)
Controls	No	Yes	No	Yes
Mean of dep. var	0.925	0.925	0.879	0.879
N	2,039	2,036	2,039	2,036

Notes: Control variables include region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A3: Selection into partner's presence (ITT)

Dependent Variable	Partner Present in HH at Endline	
Treatment	-0.00350	-0.000897
	(0.0228)	(0.0219)
No. children under age 5 in HH (baseline)	0.00430	-0.00480
	(0.00962)	(0.00945)
Respondent Age	0.00230***	0.00110
	(0.000818)	(0.000836)
Respondent literate (baseline)	0.00490	-0.000811
	(0.0236)	(0.0227)
Partner present at baseline		0.138***
		(0.0348)
Partnered at baseline		0.0604
		(0.0567)
Control mean of dep. var	0.852	0.852
N	1,337	1,337

Notes: Sample includes respondents partnered (married or cohabitating) at endline. Standard errors clustered at the level of treatment (group).

Table A4: Treatment effects on spouse's earnings (ITT)

Dependent Variable	Partner Works	Partner's Earnings	Earnings non-missing
Treatment	0.0440	0.718***	-0.00684
	(0.0324)	(0.228)	(0.0153)
Control mean of dep. var	0.314	1.175	0.941
N	1,133	1,068	1,139

Notes: Sample includes respondents partnered (married or cohabitating) at endline. Spouse's earnings winsorized at the 5th and 95th percentiles. Control variables include region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A5: Effects on Time Use Outcomes (ITT)

Dependent Variable	Minutes of time yesterday			
	Market labor	HH work	Leisure	Sleep
Panel A: Endline				
Treatment	16.25	4.189	-11.13	4.952
	(11.27)	(8.382)	(10.27)	(6.744)
	[1]	[1]	[1]	[1]
Control mean of dep. var	144.1	277.7	191.8	542.2
N	1,887	1,887	1,887	1,887
Panel B: 1 Year Follow-Up				
Treatment	-10.39	5.231	-7.745	1.701
	(11.34)	(7.531)	(9.952)	(5.392)
	[1]	[1]	[1]	[1]
P-value: $\beta^{End} = \beta^{1YFU}$	0.189	0.976	0.690	0.771
Control mean of dep. var	196.6	317.7	240.8	608.5
N	1,793	1,793	1,793	1,793

Notes: Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A6: Effects on Specific Decisions (ITT)

Dependent Variable	Respondent participates in decisions about			
	Work	HH Purchases	Medical	Med–children
Panel A: Endline				
Treatment	-0.00248	-0.00130	0.0270	0.00943
	(0.0267)	(0.0233)	(0.0264)	(0.0249)
	[1]	[1]	[1]	[1]
Control mean of dep. var	0.704	0.634	0.609	0.681
N	1,887	1,887	1,887	1,887
Panel B: 1 Year Follow-Up				
Treatment	0.0668***	0.0418*	0.0447*	0.0341
	(0.0193)	(0.0214)	(0.0269)	(0.0236)
	[.004]	[.089]	[.111]	[.123]
P-value: $\beta^{End} = \beta^{1YFU}$	0.021	0.145	0.616	0.518
Control mean of dep. var	0.719	0.641	0.576	0.641
N	1,793	1,793	1,793	1,793

Notes: Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A7: Treatment Effect Heterogeneity (ITT)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline									
				Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
Non-food consumption (USD)	Endline	Treatment	β	0.613	0.658	1.064	0.763	0.811	0.803	0.193	0.555	0.829	0.747
			SE	0.223	0.213	0.312	0.272	0.256	0.287	0.559	0.246	0.334	0.200
		Treatment x INT	β	0.665	0.295	-0.417	0.044	-0.060	-0.040	0.692	0.508	-0.292	0.236
			SE	0.411	0.355	0.351	0.365	0.378	0.404	0.589	0.329	0.482	0.475
	1YFU	Treatment	β	0.527	0.774	0.837	0.467	0.429	0.371	0.650	0.555	0.245	0.684
			SE	0.208	0.248	0.329	0.235	0.261	0.267	0.524	0.252	0.382	0.188
		Treatment x INT	β	0.324	-0.445	-0.369	0.234	0.320	0.440	0.006	0.089	0.278	-0.553
			SE	0.429	0.389	0.442	0.350	0.313	0.328	0.568	0.362	0.491	0.467
Food consumption (USD)	Endline	Treatment	β	0.786	0.820	1.554	1.103	1.140	1.151	1.007	1.093	1.412	1.106
			SE	0.356	0.361	0.464	0.408	0.413	0.441	0.726	0.430	0.463	0.323
		Treatment x INT	β	1.214	0.693	-0.644	-0.056	-0.054	-0.063	0.100	0.197	-0.537	0.024
			SE	0.594	0.539	0.559	0.444	0.473	0.465	0.764	0.521	0.554	0.612
	1YFU	Treatment	β	0.292	0.253	0.620	0.709	0.267	0.351	-0.046	0.576	0.574	0.327
			SE	0.230	0.293	0.390	0.263	0.254	0.322	0.506	0.284	0.435	0.225
		Treatment x INT	β	0.476	0.364	-0.332	-0.678	0.262	0.095	0.522	-0.240	-0.344	0.470
			SE	0.469	0.463	0.493	0.319	0.315	0.416	0.526	0.364	0.577	0.504
Women's clothes (USD)	Endline	Treatment	β	0.228	0.177	0.265	0.175	0.210	0.171	0.164	0.190	0.225	0.209
			SE	0.036	0.039	0.054	0.047	0.040	0.041	0.078	0.041	0.053	0.033
		Treatment x INT	β	-0.056	0.083	-0.079	0.066	-0.001	0.079	0.049	0.056	-0.068	0.014
			SE	0.060	0.059	0.064	0.060	0.056	0.057	0.084	0.052	0.081	0.072
	1YFU	Treatment	β	0.151	0.126	0.203	0.151	0.143	0.142	0.192	0.166	0.159	0.145
			SE	0.031	0.037	0.053	0.041	0.042	0.044	0.061	0.039	0.049	0.030
		Treatment x INT	β	0.020	0.068	-0.070	-0.009	0.021	0.027	-0.039	-0.034	-0.071	0.058
			SE	0.062	0.055	0.063	0.048	0.052	0.059	0.069	0.054	0.071	0.063
Men's clothes (USD)	Endline	Treatment	β	0.039	0.028	0.013	0.035	0.043	0.018	0.029	0.037	0.038	0.032
			SE	0.011	0.013	0.013	0.013	0.013	0.013	0.032	0.012	0.016	0.011
		Treatment x INT	β	-0.003	0.022	0.035	0.003	-0.012	0.038	0.009	0.002	0.011	0.037
			SE	0.023	0.021	0.018	0.019	0.017	0.018	0.034	0.016	0.026	0.027
	1YFU	Treatment	β	0.037	0.023	0.029	0.035	0.042	0.029	0.040	0.035	0.049	0.026
			SE	0.012	0.014	0.016	0.015	0.016	0.016	0.025	0.014	0.022	0.011
		Treatment x INT	β	-0.030	0.009	-0.003	-0.015	-0.028	-0.001	-0.011	-0.016	-0.036	0.012
			SE	0.019	0.022	0.023	0.019	0.019	0.020	0.027	0.017	0.028	0.027
Girls' clothes (USD)	Endline	Treatment	β	0.039	0.042	0.063	0.044	0.040	0.041	0.027	0.031	0.042	0.044
			SE	0.008	0.011	0.014	0.012	0.011	0.011	0.025	0.010	0.018	0.009
		Treatment x INT	β	0.018	0.007	-0.028	-0.002	0.006	0.005	0.018	0.025	-0.015	-0.001
			SE	0.020	0.019	0.019	0.019	0.018	0.016	0.026	0.017	0.025	0.024
	1YFU	Treatment	β	0.042	0.027	0.016	0.029	0.028	0.032	0.062	0.039	0.041	0.031
			SE	0.011	0.012	0.016	0.015	0.015	0.014	0.027	0.016	0.018	0.010
		Treatment x INT	β	-0.021	0.018	0.029	0.012	0.015	0.006	-0.029	-0.008	-0.006	0.029
			SE	0.022	0.018	0.020	0.019	0.019	0.019	0.030	0.021	0.025	0.024

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity (continued)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline									
				Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
Boys' clothes (USD)	Endline	Treatment	β	0.036	0.035	0.037	0.027	0.033	0.029	0.035	0.026	0.042	0.040
			SE	0.007	0.009	0.012	0.008	0.008	0.009	0.018	0.007	0.012	0.007
		Treatment x INT	β	0.006	0.007	0.001	0.020	0.010	0.017	0.001	0.016	-0.007	-0.014
			SE	0.014	0.013	0.013	0.013	0.011	0.013	0.019	0.012	0.018	0.014
	1YFU	Treatment	β	0.017	0.002	-0.001	-0.003	0.000	0.001	0.016	0.013	0.012	0.006
			SE	0.009	0.010	0.012	0.011	0.011	0.011	0.016	0.012	0.016	0.008
		Treatment x INT	β	-0.027	0.016	0.013	0.023	0.016	0.016	-0.008	-0.008	0.001	0.016
			SE	0.017	0.014	0.016	0.016	0.015	0.015	0.018	0.014	0.024	0.019
Assets (standardized)	Endline	Treatment	β	0.145	0.148	0.238	0.190	0.218	0.204	0.322	0.164	0.232	0.186
			SE	0.046	0.063	0.063	0.061	0.068	0.061	0.118	0.065	0.089	0.049
		Treatment x INT	β	0.084	0.034	-0.108	-0.062	-0.104	-0.074	-0.172	0.041	-0.210	-0.139
			SE	0.092	0.098	0.071	0.088	0.086	0.085	0.117	0.077	0.104	0.114
	1YFU	Treatment	β	0.095	0.034	0.104	0.047	0.060	0.086	0.069	0.091	0.029	0.073
			SE	0.042	0.058	0.078	0.063	0.064	0.057	0.104	0.059	0.088	0.045
		Treatment x INT	β	-0.057	0.086	-0.044	0.043	0.024	-0.025	0.004	-0.019	0.062	-0.001
			SE	0.092	0.087	0.093	0.084	0.084	0.082	0.110	0.080	0.104	0.085
Livestock (standardized)	Endline	Treatment	β	0.113	0.142	0.157	0.048	0.077	0.079	0.165	0.089	0.162	0.156
			SE	0.050	0.026	0.035	0.072	0.069	0.069	0.051	0.065	0.047	0.023
		Treatment x INT	β	0.036	-0.036	-0.048	0.145	0.091	0.088	-0.049	0.085	-0.105	-0.195
			SE	0.059	0.077	0.065	0.073	0.071	0.070	0.071	0.068	0.104	0.205
	1YFU	Treatment	β	0.212	0.221	0.281	0.180	0.155	0.108	0.219	0.188	0.315	0.239
			SE	0.061	0.062	0.080	0.058	0.062	0.062	0.164	0.060	0.074	0.045
		Treatment x INT	β	0.054	0.018	-0.080	0.079	0.134	0.233	0.001	0.058	-0.217	-0.070
			SE	0.106	0.083	0.095	0.086	0.072	0.093	0.177	0.098	0.101	0.097
Total earnings (USD)	Endline	Treatment	β	0.135	0.118	0.289	0.241	0.316	0.189	0.357	0.064	-0.063	0.193
			SE	0.110	0.110	0.159	0.131	0.123	0.144	0.270	0.136	0.191	0.103
		Treatment x INT	β	0.200	0.197	-0.138	-0.110	-0.241	0.014	-0.192	0.216	0.347	-0.002
			SE	0.230	0.175	0.165	0.148	0.166	0.203	0.282	0.165	0.244	0.235
	1YFU	Treatment	β	0.434	0.604	0.362	0.560	0.506	0.560	0.748	0.589	0.450	0.527
			SE	0.144	0.139	0.201	0.151	0.170	0.144	0.282	0.179	0.205	0.131
		Treatment x INT	β	0.171	-0.288	0.161	-0.186	-0.056	-0.166	-0.324	-0.229	0.140	-0.337
			SE	0.276	0.216	0.246	0.203	0.200	0.216	0.309	0.224	0.305	0.327
Earnings net of costs (USD)	Endline	Treatment	β	0.022	0.018	0.173	0.114	0.166	0.000	0.061	-0.014	-0.053	0.054
			SE	0.074	0.083	0.112	0.085	0.088	0.100	0.182	0.082	0.139	0.074
		Treatment x INT	β	0.130	0.137	-0.161	-0.096	-0.204	0.126	-0.003	0.128	0.133	0.032
			SE	0.147	0.135	0.125	0.120	0.119	0.137	0.203	0.119	0.161	0.165
	1YFU	Treatment	β	0.199	0.261	0.190	0.271	0.268	0.322	0.323	0.360	0.162	0.235
			SE	0.098	0.085	0.149	0.110	0.107	0.094	0.188	0.118	0.133	0.082
		Treatment x INT	β	-0.005	-0.117	0.006	-0.152	-0.134	-0.245	-0.161	-0.329	0.046	-0.252
			SE	0.190	0.149	0.177	0.152	0.145	0.150	0.214	0.153	0.192	0.216

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity (continued)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline									
				Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
Total business costs (USD)	Endline	Treatment	β	0.188	0.155	0.256	0.171	0.150	0.222	0.313	0.163	-0.112	0.177
			SE	0.078	0.081	0.112	0.104	0.090	0.086	0.192	0.106	0.133	0.072
		Treatment x INT	β	-0.015	0.066	-0.111	-0.012	0.065	-0.080	-0.154	0.013	0.385	0.025
			SE	0.159	0.128	0.132	0.113	0.134	0.123	0.204	0.136	0.176	0.175
	1YFU	Treatment	β	0.292	0.380	0.232	0.354	0.276	0.289	0.409	0.235	0.318	0.312
			SE	0.096	0.111	0.136	0.109	0.120	0.121	0.238	0.114	0.155	0.092
		Treatment x INT	β	0.130	-0.142	0.133	-0.066	0.097	0.067	-0.091	0.170	0.109	0.076
			SE	0.161	0.145	0.159	0.141	0.153	0.154	0.230	0.152	0.196	0.219
Worked last week	Endline	Treatment	β	-0.022	-0.035	0.009	-0.004	0.027	-0.009	0.014	-0.014	-0.104	-0.004
			SE	0.026	0.029	0.037	0.030	0.033	0.033	0.061	0.030	0.047	0.023
		Treatment x INT	β	0.054	0.074	-0.020	-0.007	-0.063	0.009	-0.021	-0.006	0.136	-0.004
			SE	0.052	0.044	0.049	0.044	0.044	0.046	0.066	0.049	0.064	0.058
	1YFU	Treatment	β	0.043	0.046	0.039	0.070	0.038	0.071	0.120	0.044	0.025	0.051
			SE	0.033	0.029	0.044	0.036	0.035	0.031	0.057	0.029	0.042	0.027
		Treatment x INT	β	-0.002	-0.008	0.006	-0.049	0.010	-0.054	-0.091	-0.010	0.024	-0.058
			SE	0.060	0.043	0.054	0.051	0.049	0.048	0.066	0.043	0.058	0.061
Hours of work last week	Endline	Treatment	β	-0.273	0.256	0.055	-0.259	0.010	0.449	0.815	-0.024	-2.233	-0.223
			SE	0.905	0.872	1.400	1.193	0.978	0.989	1.996	1.168	1.402	0.800
		Treatment x INT	β	0.739	-0.302	-0.091	0.410	0.000	-0.882	-0.880	-0.339	4.115	1.363
			SE	1.504	1.397	1.452	1.458	1.225	1.390	2.191	1.234	1.771	1.910
	1YFU	Treatment	β	2.060	2.144	2.756	1.574	0.723	2.001	2.475	1.645	-0.079	2.039
			SE	0.881	0.995	1.385	0.942	0.956	0.943	1.549	1.028	1.241	0.867
		Treatment x INT	β	-1.864	-1.276	-1.808	-0.146	1.567	-0.984	-1.228	-0.673	2.067	-3.510
			SE	1.776	1.540	1.387	1.249	1.258	1.341	1.651	1.446	1.596	2.011
Is self employed	Endline	Treatment	β	0.066	0.059	0.066	0.041	0.058	0.063	0.054	0.042	0.021	0.052
			SE	0.021	0.023	0.032	0.026	0.025	0.023	0.047	0.028	0.030	0.019
		Treatment x INT	β	-0.012	0.009	-0.006	0.034	0.009	-0.001	0.011	0.033	0.063	0.065
			SE	0.040	0.036	0.038	0.035	0.036	0.030	0.050	0.033	0.043	0.043
	1YFU	Treatment	β	0.073	0.072	0.053	0.066	0.057	0.069	0.059	0.069	0.036	0.069
			SE	0.021	0.022	0.026	0.025	0.026	0.023	0.047	0.027	0.040	0.019
		Treatment x INT	β	-0.013	-0.010	0.022	0.003	0.022	-0.002	0.011	0.001	0.068	-0.002
			SE	0.048	0.030	0.032	0.033	0.030	0.028	0.050	0.035	0.051	0.050
Own savings	Endline	Treatment	β	0.543	0.491	0.575	0.510	0.528	0.510	0.599	0.482	0.559	0.519
			SE	0.029	0.033	0.036	0.031	0.035	0.035	0.053	0.034	0.041	0.028
		Treatment x INT	β	-0.077	0.072	-0.080	0.016	-0.017	0.020	-0.096	0.071	-0.111	0.009
			SE	0.044	0.042	0.046	0.038	0.038	0.035	0.060	0.039	0.051	0.055
	1YFU	Treatment	β	0.366	0.337	0.326	0.360	0.352	0.374	0.428	0.367	0.383	0.366
			SE	0.034	0.038	0.044	0.037	0.040	0.038	0.055	0.037	0.048	0.032
		Treatment x INT	β	0.009	0.072	0.058	0.006	0.029	-0.013	-0.074	-0.001	-0.004	0.012
			SE	0.047	0.043	0.052	0.044	0.044	0.042	0.058	0.038	0.053	0.052

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity (continued)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline: Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
Savings (USD)	Endline	Treatment	β	6.344	6.512	7.689	5.993	6.169	6.187	7.891	6.103	7.018	6.545
			SE	0.639	0.689	0.942	0.670	0.719	0.703	1.374	0.724	0.812	0.625
		Treatment x INT	β	0.845	0.044	-1.670	0.867	0.750	0.710	-1.552	0.755	-1.864	-0.023
			SE	0.889	0.957	0.942	0.723	0.756	0.815	1.395	0.805	0.993	1.119
	1YFU	Treatment	β	7.830	8.172	8.031	7.977	7.746	8.159	7.147	8.732	7.071	8.482
			SE	0.831	0.936	1.030	0.897	1.011	0.948	1.583	0.992	1.126	0.760
		Treatment x INT	β	1.513	-0.028	0.201	0.060	0.856	0.080	1.230	-1.000	1.874	-1.808
			SE	1.507	1.193	1.199	1.092	1.183	1.107	1.588	1.162	1.569	1.319
VSLA member	Endline	Treatment	β	0.627	0.564	0.629	0.584	0.600	0.612	0.657	0.592	0.618	0.604
			SE	0.031	0.031	0.034	0.031	0.036	0.032	0.050	0.037	0.042	0.028
		Treatment x INT	β	-0.075	0.100	-0.037	0.039	0.009	-0.014	-0.063	0.016	-0.058	0.008
			SE	0.044	0.034	0.038	0.030	0.040	0.034	0.050	0.038	0.047	0.038
	1YFU	Treatment	β	0.395	0.409	0.415	0.405	0.387	0.406	0.438	0.412	0.465	0.417
			SE	0.039	0.039	0.041	0.038	0.041	0.041	0.058	0.038	0.042	0.034
		Treatment x INT	β	0.046	-0.003	-0.013	0.003	0.040	0.003	-0.037	-0.009	-0.100	-0.061
			SE	0.054	0.041	0.045	0.041	0.047	0.038	0.063	0.035	0.046	0.056
Risk tolerance	Endline	Treatment	β	0.049	0.014	-0.007	0.173	0.083	0.136	0.094	0.074	0.044	0.039
			SE	0.108	0.110	0.153	0.129	0.123	0.123	0.231	0.132	0.146	0.099
		Treatment x INT	β	0.163	0.241	0.153	-0.175	0.023	-0.074	0.013	0.084	0.107	0.395
			SE	0.191	0.164	0.174	0.173	0.144	0.166	0.244	0.151	0.187	0.233
	1YFU	Treatment	β	0.300	0.322	0.218	0.341	0.375	0.243	0.470	0.396	0.205	0.293
			SE	0.122	0.114	0.144	0.126	0.121	0.128	0.211	0.123	0.146	0.113
		Treatment x INT	β	-0.025	-0.059	0.104	-0.120	-0.175	0.091	-0.195	-0.179	0.205	-0.036
			SE	0.188	0.150	0.148	0.169	0.152	0.167	0.223	0.159	0.201	0.222
Partnered	Endline	Treatment	β	0.008	-0.029	-0.000	0.011	-0.014	-0.025	-0.033	-0.014	0.002	-0.001
			SE	0.018	0.018	0.034	0.022	0.022	0.019	0.043	0.024	0.024	0.017
		Treatment x INT	β	-0.044	0.059	-0.006	-0.021	0.019	0.041	0.034	0.033	-0.004	-0.016
			SE	0.034	0.030	0.037	0.033	0.031	0.026	0.048	0.032	0.026	0.039
	1YFU	Treatment	β	-0.008	-0.020	0.040	0.017	-0.002	-0.025	0.030	-0.032	-0.004	0.002
			SE	0.019	0.024	0.039	0.023	0.025	0.027	0.042	0.024	0.024	0.020
		Treatment x INT	β	0.014	0.035	-0.065	-0.036	-0.006	0.041	-0.038	0.069	-0.035	-0.041
			SE	0.040	0.035	0.039	0.031	0.032	0.031	0.044	0.034	0.032	0.037
Pro-women attitudes index	Endline	Treatment	β	0.200	0.088	0.039	0.178	0.126	0.277	0.220	0.181	0.373	0.128
			SE	0.064	0.062	0.080	0.068	0.068	0.070	0.124	0.073	0.074	0.057
		Treatment x INT	β	-0.028	0.256	0.216	0.020	0.116	-0.177	-0.034	0.033	-0.253	0.388
			SE	0.103	0.094	0.090	0.091	0.096	0.090	0.135	0.102	0.108	0.123
	1YFU	Treatment	β	0.147	0.177	0.121	0.196	0.206	0.213	0.263	0.200	0.155	0.113
			SE	0.108	0.109	0.145	0.123	0.139	0.122	0.197	0.125	0.165	0.099
		Treatment x INT	β	-0.080	-0.092	0.010	-0.172	-0.160	-0.171	-0.162	-0.123	0.035	0.102
			SE	0.179	0.171	0.156	0.148	0.164	0.159	0.217	0.161	0.201	0.218

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity (continued)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline									
				Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
IPV index	Endline	Treatment	β	-0.184	-0.214	-0.246	-0.092	-0.068	-0.129	0.253	0.526	-0.019	-0.172
			SE	0.112	0.121	0.291	0.120	0.115	0.113	0.252	0.117	0.162	0.111
		Treatment x INT	β	0.293	0.293	0.164	0.005	-0.065	0.057	-0.421	-1.686	-0.090	0.336
			SE	0.165	0.168	0.307	0.159	0.149	0.137	0.284	0.188	0.214	0.182
	1YFU	Treatment	β	-0.101	-0.058	-0.098	0.091	0.049	-0.156	-0.149	0.637	0.097	-0.013
			SE	0.100	0.099	0.174	0.114	0.109	0.107	0.256	0.108	0.166	0.096
		Treatment x INT	β	0.327	0.122	0.106	-0.183	-0.113	0.297	0.148	-1.770	-0.191	0.018
			SE	0.170	0.168	0.212	0.180	0.172	0.159	0.282	0.190	0.247	0.207
Decisions index	Endline	Treatment	β	-0.002	-0.033	-0.006	0.064	0.084	0.054	0.229	-0.045	0.140	-0.004
			SE	0.059	0.065	0.084	0.073	0.075	0.084	0.142	0.067	0.088	0.059
		Treatment x INT	β	0.078	0.136	0.038	-0.090	-0.126	-0.067	-0.241	0.163	-0.131	0.172
			SE	0.091	0.094	0.098	0.083	0.096	0.094	0.143	0.083	0.103	0.121
	1YFU	Treatment	β	0.207	0.183	0.034	0.188	0.188	0.100	0.138	0.094	0.316	0.164
			SE	0.067	0.077	0.092	0.081	0.086	0.087	0.169	0.084	0.118	0.061
		Treatment x INT	β	-0.151	-0.060	0.187	-0.071	-0.044	0.124	0.038	0.159	-0.106	-0.022
			SE	0.119	0.120	0.118	0.111	0.118	0.117	0.183	0.110	0.144	0.149
Locus of control	Endline	Treatment	β	0.131	-0.002	0.006	0.046	0.084	0.101	-0.124	0.081	-0.004	0.037
			SE	0.060	0.069	0.087	0.065	0.073	0.071	0.127	0.073	0.091	0.057
		Treatment x INT	β	-0.297	0.119	0.059	-0.023	-0.074	-0.106	0.205	-0.034	0.082	0.066
			SE	0.096	0.103	0.093	0.086	0.095	0.092	0.128	0.097	0.116	0.128
	1YFU	Treatment	β	0.165	0.167	0.078	0.133	0.137	0.142	0.271	0.206	0.084	0.139
			SE	0.059	0.065	0.092	0.065	0.063	0.068	0.132	0.067	0.099	0.057
		Treatment x INT	β	-0.076	-0.054	0.093	0.030	0.013	0.003	-0.156	-0.112	0.178	0.029
			SE	0.102	0.092	0.104	0.097	0.101	0.087	0.147	0.100	0.131	0.125
Depression index (PHQ-9)	Endline	Treatment	β	-0.051	0.060	0.130	0.126	0.174	0.093	1.112	0.265	0.017	-0.128
			SE	0.288	0.319	0.394	0.323	0.340	0.318	0.497	0.331	0.464	0.261
		Treatment x INT	β	-0.492	-0.588	-0.440	-0.548	-0.644	-0.525	-1.471	-1.104	-0.383	-0.285
			SE	0.469	0.456	0.448	0.434	0.451	0.366	0.546	0.430	0.532	0.540
	1YFU	Treatment	β	-0.254	-0.259	-1.255	-0.410	-0.240	-0.410	-0.038	-0.146	0.449	-0.386
			SE	0.314	0.339	0.418	0.324	0.323	0.374	0.724	0.341	0.443	0.279
		Treatment x INT	β	-0.212	-0.075	1.395	0.267	-0.111	0.227	-0.339	-0.489	-0.144	0.593
			SE	0.514	0.469	0.509	0.461	0.377	0.523	0.785	0.483	0.521	0.573
Anxiety index (GAD7)	Endline	Treatment	β	-0.434	-0.228	-0.025	0.208	0.210	-0.132	0.877	0.435	-0.070	-0.191
			SE	0.335	0.354	0.449	0.357	0.330	0.350	0.643	0.350	0.432	0.301
		Treatment x INT	β	0.465	-0.131	-0.387	-0.920	-0.998	-0.316	-1.340	-1.663	-0.243	-0.603
			SE	0.488	0.502	0.476	0.470	0.433	0.421	0.681	0.452	0.601	0.531
	1YFU	Treatment	β	-0.416	-0.334	-1.063	-0.398	-0.139	-0.751	0.147	-0.358	0.084	-0.534
			SE	0.335	0.328	0.426	0.367	0.348	0.329	0.715	0.322	0.463	0.298
		Treatment x INT	β	-0.206	-0.204	0.892	-0.055	-0.632	0.584	-0.730	-0.273	0.050	0.531
			SE	0.549	0.471	0.438	0.527	0.462	0.438	0.770	0.486	0.605	0.636

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity (continued)

Dependent Variable	Timing	Coeff.		Interaction with [...], as defined at baseline									
				Lit	Work	Part	Depr	Anxious	Dec	Atti	Pred IPV	Age Diff	Main Earner
Physical health index (ADLs)	Endline	Treatment	β	-0.017	-0.035	0.035	-0.079	-0.045	-0.033	-0.099	-0.003	-0.057	-0.008
			SE	0.066	0.071	0.081	0.070	0.075	0.079	0.130	0.076	0.112	0.061
		Treatment x INT	β	0.078	0.111	-0.044	0.139	0.094	0.073	0.115	0.010	0.068	0.083
			SE	0.113	0.098	0.090	0.091	0.093	0.085	0.144	0.112	0.139	0.123
	1YFU	Treatment	β	0.101	0.093	0.374	0.222	0.142	0.134	-0.187	0.147	0.084	0.157
			SE	0.108	0.095	0.143	0.131	0.115	0.103	0.198	0.118	0.176	0.089
		Treatment x INT	β	0.111	0.087	-0.355	-0.226	-0.023	-0.012	0.344	-0.034	-0.056	-0.193
			SE	0.157	0.143	0.154	0.161	0.158	0.145	0.214	0.167	0.214	0.184

Notes: Each coefficient is from a separate regression. Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Asset, livestock, pro-women attitudes, IPV, decisions, and physical health indices constructed following (Anderson, 2008) and standardized so that the control group at endline has mean zero and standard deviation one.Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A8: Treatment Effect Heterogeneity: Indices (ITT)

Dependent Variable	Timing	Coeff.		Above median value for: SES	Health	Social	Inv Pred IPV
Non-food consumption (USD)	Endline	Treatment	β	0.664	1.031	0.908	0.555
			SE	0.262	0.265	0.258	0.246
		Treatment x INT	β	0.228	-0.507	-0.252	0.508
			SE	0.327	0.371	0.328	0.329
	1YFU	Treatment	β	0.445	0.327	0.410	0.555
			SE	0.280	0.290	0.259	0.252
		Treatment x INT	β	0.280	0.529	0.361	0.089
			SE	0.390	0.364	0.358	0.362
Food consumption (USD)	Endline	Treatment	β	0.966	1.279	1.025	1.093
			SE	0.389	0.379	0.401	0.430
		Treatment x INT	β	0.298	-0.326	0.183	0.197
			SE	0.457	0.501	0.508	0.521
	1YFU	Treatment	β	0.071	0.071	0.494	0.576
			SE	0.305	0.284	0.284	0.284
		Treatment x INT	β	0.640	0.648	-0.195	-0.240
			SE	0.415	0.334	0.443	0.364
Women's clothes (USD)	Endline	Treatment	β	0.222	0.186	0.262	0.190
			SE	0.041	0.042	0.039	0.041
		Treatment x INT	β	-0.023	0.050	-0.102	0.056
			SE	0.052	0.054	0.048	0.052
	1YFU	Treatment	β	0.132	0.124	0.144	0.166
			SE	0.042	0.035	0.040	0.039
		Treatment x INT	β	0.043	0.062	0.020	-0.034
			SE	0.055	0.050	0.053	0.054
Men's clothes (USD)	Endline	Treatment	β	0.034	0.038	0.051	0.037
			SE	0.013	0.014	0.012	0.012
		Treatment x INT	β	0.007	-0.002	-0.028	0.002
			SE	0.018	0.018	0.016	0.016
	1YFU	Treatment	β	0.024	0.014	0.015	0.035
			SE	0.016	0.014	0.013	0.014
		Treatment x INT	β	0.007	0.028	0.026	-0.016
			SE	0.019	0.016	0.019	0.017
Girls' clothes (USD)	Endline	Treatment	β	0.038	0.047	0.051	0.031
			SE	0.010	0.010	0.009	0.010
		Treatment x INT	β	0.012	-0.006	-0.015	0.025
			SE	0.015	0.015	0.013	0.017
	1YFU	Treatment	β	0.028	0.034	0.044	0.039
			SE	0.014	0.014	0.012	0.016
		Treatment x INT	β	0.014	0.004	-0.016	-0.008
			SE	0.021	0.020	0.021	0.021
Boys' clothes (USD)	Endline	Treatment	β	0.039	0.039	0.040	0.026
			SE	0.009	0.009	0.010	0.007
		Treatment x INT	β	-0.003	-0.002	-0.005	0.016
			SE	0.012	0.012	0.012	0.012
	1YFU	Treatment	β	-0.003	0.011	0.020	0.013
			SE	0.012	0.010	0.010	0.012
		Treatment x INT	β	0.022	-0.005	-0.022	-0.008
			SE	0.015	0.014	0.016	0.014
Assets (standardized)	Endline	Treatment	β	0.166	0.077	0.133	0.164
			SE	0.054	0.062	0.066	0.065
		Treatment x INT	β	-0.002	0.177	0.064	0.041
			SE	0.072	0.081	0.087	0.077
	1YFU	Treatment	β	0.111	0.031	0.031	0.091
			SE	0.047	0.059	0.058	0.059
		Treatment x INT	β	-0.075	0.086	0.085	-0.019
			SE	0.075	0.078	0.072	0.080

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity: Indices (continued)

Dependent Variable	Timing	Coeff.		Above median value for: SES	Health	Social	Inv Pred IPV
Livestock (standardized)	Endline	Treatment	β	0.106	0.077	0.098	0.089
			SE	0.071	0.067	0.069	0.065
		Treatment x INT	β	0.036	0.096	0.052	0.085
			SE	0.079	0.074	0.074	0.068
	1YFU	Treatment	β	0.259	0.262	0.258	0.188
			SE	0.067	0.074	0.053	0.060
		Treatment x INT	β	-0.065	-0.070	-0.063	0.058
			SE	0.086	0.097	0.079	0.098
Total earnings (USD)	Endline	Treatment	β	0.233	0.247	0.198	0.064
			SE	0.150	0.132	0.116	0.136
		Treatment x INT	β	-0.084	-0.105	-0.011	0.216
			SE	0.184	0.168	0.160	0.165
	1YFU	Treatment	β	0.600	0.709	0.399	0.589
			SE	0.162	0.120	0.136	0.179
		Treatment x INT	β	-0.251	-0.470	0.148	-0.229
			SE	0.216	0.180	0.202	0.224
Earnings net of costs (USD)	Endline	Treatment	β	0.085	0.048	-0.022	-0.014
			SE	0.102	0.102	0.081	0.082
		Treatment x INT	β	-0.047	0.031	0.167	0.128
			SE	0.134	0.140	0.115	0.119
	1YFU	Treatment	β	0.284	0.281	0.123	0.360
			SE	0.116	0.087	0.096	0.118
		Treatment x INT	β	-0.177	-0.172	0.146	-0.329
			SE	0.161	0.135	0.132	0.153
Total business costs (USD)	Endline	Treatment	β	0.244	0.237	0.234	0.163
			SE	0.101	0.091	0.085	0.106
		Treatment x INT	β	-0.131	-0.114	-0.110	0.013
			SE	0.120	0.115	0.126	0.136
	1YFU	Treatment	β	0.382	0.498	0.299	0.235
			SE	0.132	0.111	0.101	0.114
		Treatment x INT	β	-0.119	-0.351	0.044	0.170
			SE	0.171	0.133	0.135	0.152
Worked last week	Endline	Treatment	β	0.001	0.008	-0.013	-0.014
			SE	0.033	0.034	0.032	0.030
		Treatment x INT	β	-0.013	-0.027	0.017	-0.006
			SE	0.042	0.046	0.048	0.049
	1YFU	Treatment	β	0.059	0.065	0.009	0.044
			SE	0.035	0.029	0.036	0.029
		Treatment x INT	β	-0.034	-0.046	0.068	-0.010
			SE	0.049	0.040	0.044	0.043
Hours of work last week	Endline	Treatment	β	0.432	0.080	0.050	-0.024
			SE	1.159	0.938	1.081	1.168
		Treatment x INT	β	-0.911	-0.199	-0.131	-0.339
			SE	1.291	1.414	1.297	1.234
	1YFU	Treatment	β	2.265	2.141	1.288	1.645
			SE	1.069	0.951	0.850	1.028
		Treatment x INT	β	-1.554	-1.296	0.397	-0.673
			SE	1.262	1.292	1.260	1.446
Is self employed	Endline	Treatment	β	0.071	0.070	0.059	0.042
			SE	0.026	0.021	0.025	0.028
		Treatment x INT	β	-0.018	-0.017	0.005	0.033
			SE	0.028	0.034	0.031	0.033
	1YFU	Treatment	β	0.100	0.086	0.040	0.069
			SE	0.022	0.023	0.024	0.027
		Treatment x INT	β	-0.062	-0.037	0.055	0.001
			SE	0.030	0.035	0.031	0.035

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity: Indices (continued)

Dependent Variable	Timing	Coeff.		Above median value for			
				SES	Health	Social	Inv Pred IPV
Own savings	Endline	Treatment	β	0.529	0.546	0.599	0.482
			SE	0.038	0.027	0.031	0.034
		Treatment x INT	β	-0.018	-0.052	-0.156	0.071
			SE	0.041	0.034	0.046	0.039
	1YFU	Treatment	β	0.336	0.367	0.383	0.367
			SE	0.037	0.035	0.039	0.037
		Treatment x INT	β	0.062	-0.000	-0.033	-0.001
			SE	0.046	0.039	0.046	0.038
Savings (USD)	Endline	Treatment	β	7.062	6.104	6.998	6.103
			SE	0.843	0.629	0.629	0.724
		Treatment x INT	β	-1.054	0.877	-0.916	0.755
			SE	0.868	0.722	0.797	0.805
	1YFU	Treatment	β	8.314	8.038	7.924	8.732
			SE	1.010	0.940	0.842	0.992
		Treatment x INT	β	-0.236	0.322	0.521	-1.000
			SE	1.165	0.924	1.092	1.162
VSLA member	Endline	Treatment	β	0.615	0.620	0.663	0.592
			SE	0.034	0.028	0.030	0.037
		Treatment x INT	β	-0.020	-0.031	-0.117	0.016
			SE	0.035	0.030	0.038	0.038
	1YFU	Treatment	β	0.392	0.384	0.449	0.412
			SE	0.038	0.040	0.039	0.038
		Treatment x INT	β	0.030	0.046	-0.084	-0.009
			SE	0.041	0.040	0.038	0.035
Risk tolerance	Endline	Treatment	β	0.165	0.204	0.025	0.074
			SE	0.120	0.130	0.117	0.132
		Treatment x INT	β	-0.143	-0.218	0.142	0.084
			SE	0.147	0.149	0.163	0.151
	1YFU	Treatment	β	0.227	0.066	0.123	0.396
			SE	0.142	0.123	0.130	0.123
		Treatment x INT	β	0.124	0.449	0.333	-0.179
			SE	0.159	0.134	0.150	0.159
Partnered	Endline	Treatment	β	0.002	0.006	0.019	-0.014
			SE	0.024	0.021	0.023	0.024
		Treatment x INT	β	-0.013	-0.020	-0.046	0.033
			SE	0.029	0.027	0.032	0.032
	1YFU	Treatment	β	0.005	-0.034	-0.001	-0.032
			SE	0.026	0.025	0.023	0.024
		Treatment x INT	β	-0.018	0.059	-0.008	0.069
			SE	0.032	0.031	0.032	0.034
Pro-women attitudes index	Endline	Treatment	β	0.138	0.203	0.218	0.181
			SE	0.069	0.073	0.067	0.073
		Treatment x INT	β	0.098	-0.032	-0.062	0.033
			SE	0.091	0.096	0.084	0.102
	1YFU	Treatment	β	-0.095	0.129	0.214	0.200
			SE	0.115	0.115	0.131	0.125
		Treatment x INT	β	0.439	-0.003	-0.172	-0.123
			SE	0.142	0.139	0.151	0.161
IPV index	Endline	Treatment	β	-0.101	-0.309	-0.167	0.526
			SE	0.123	0.134	0.141	0.117
		Treatment x INT	β	0.002	0.426	0.128	-1.686
			SE	0.156	0.176	0.189	0.188
	1YFU	Treatment	β	-0.102	-0.081	-0.118	0.637
			SE	0.120	0.119	0.115	0.108
		Treatment x INT	β	0.178	0.142	0.210	-1.770
			SE	0.153	0.179	0.169	0.190

Notes: Business costs include the discounted use value of large purchases. Earnings, costs, and savings winsorized at the 5th and 95th percentiles. Numbers in brackets are sharpened q-values (Benjamini et al., 2006) that control the false discovery rate. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Treatment Effect Heterogeneity: Indices (continued)

Dependent Variable	Timing	Coeff.		Above median value for			
				SES	Health	Social	Inv Pred IPV
Decisions index	Endline	Treatment	β	0.035	0.076	0.040	-0.045
			SE	0.071	0.069	0.064	0.067
		Treatment x INT	β	-0.028	-0.110	-0.039	0.163
			SE	0.086	0.096	0.077	0.083
	1YFU	Treatment	β	0.196	0.126	0.135	0.094
			SE	0.088	0.083	0.084	0.084
		Treatment x INT	β	-0.065	0.072	0.055	0.159
			SE	0.120	0.099	0.121	0.110
Locus of control	Endline	Treatment	β	0.051	-0.022	0.017	0.081
			SE	0.069	0.058	0.072	0.073
		Treatment x INT	β	-0.010	0.135	0.057	-0.034
			SE	0.087	0.092	0.104	0.097
	1YFU	Treatment	β	0.170	0.094	0.151	0.206
			SE	0.066	0.073	0.067	0.067
		Treatment x INT	β	-0.054	0.096	-0.016	-0.112
			SE	0.086	0.108	0.085	0.100
Depression index (PHQ-9)	Endline	Treatment	β	-0.313	-0.303	-0.119	0.265
			SE	0.360	0.362	0.344	0.331
		Treatment x INT	β	0.286	0.267	-0.098	-1.104
			SE	0.471	0.410	0.390	0.430
	1YFU	Treatment	β	-0.033	0.272	-0.314	-0.146
			SE	0.372	0.389	0.356	0.341
		Treatment x INT	β	-0.505	-1.140	0.045	-0.489
			SE	0.475	0.471	0.481	0.483
Anxiety index (GAD7)	Endline	Treatment	β	-0.496	-0.456	-0.626	0.435
			SE	0.398	0.405	0.379	0.350
		Treatment x INT	β	0.409	0.346	0.673	-1.663
			SE	0.431	0.439	0.457	0.452
	1YFU	Treatment	β	-0.539	0.082	-0.339	-0.358
			SE	0.369	0.361	0.339	0.322
		Treatment x INT	β	0.174	-1.077	-0.222	-0.273
			SE	0.472	0.426	0.449	0.486
Physical health index (ADLs)	Endline	Treatment	β	-0.015	-0.012	-0.023	-0.003
			SE	0.069	0.067	0.077	0.076
		Treatment x INT	β	0.037	0.033	0.055	0.010
			SE	0.104	0.096	0.102	0.112
	1YFU	Treatment	β	0.107	-0.032	0.143	0.147
			SE	0.124	0.106	0.112	0.118
		Treatment x INT	β	0.044	0.331	-0.025	-0.034
			SE	0.167	0.145	0.148	0.167

Notes: Each coefficient is from a separate regression. Interaction variables refer to whether the respondent is above the median value in the indices described in section A.1, which include only baseline covariates. Control variables include the dependent variable at baseline, region dummies, a quadratic in age, and the following variables (all defined at baseline) which were unbalanced at baseline: number of children under 5 in household, work hours, decisions index, pro-women attitudes, 1(pro-women attitudes non-missing), PHQ-9 depression index, 1(PHQ-9 depression index non-missing). Standard errors clustered at the level of treatment (group).

Table A9: Correlation between indices

Dependent Variable	Inverse IPV	Economic	Health	Social
Inverse IPV	1.0000			
Economic	0.0267	1.0000		
Health	0.1879	0.0726	1.0000	
Social	0.0107	0.0068	0.0411	1.0000

Notes: Sample includes respondents partnered (married or cohabitating) at endline and 1 year follow-up. Standard errors clustered at the level of treatment (group).