



Regular article

Temporary acceleration or long-term change? Experimental evidence on female empowerment and intimate partner violence in urban Liberia[☆]David Sungho Park[✉]

KDI School of Public Policy and Management, Republic of Korea

ARTICLE INFO

JEL classification:

C93
D13
I12
J12
J16
O12
O15

Keywords:

Intimate partner violence
Psychosocial intervention
Women's empowerment
Social norms
Long-term effects
Liberia

ABSTRACT

This paper evaluates a multifaceted female empowerment program in urban Liberia that combined intensive psychosocial therapy with vocational skills training over 12 months. The evaluation uses a randomized controlled trial with follow-up surveys at 1 year and 3.5 years post-program (2 and 4.5 years after enrollment). At the first endline, treatment women experienced large reductions in emotional and physical/sexual intimate partner violence (IPV) and in perceived community norms supporting IPV; effects on psychological well-being were null. By 3.5 years, the treatment group's IPV levels remained close to their first-endline values, but the treatment–control difference had closed because the control group also experienced substantial reductions. Only the gap in perceived community norms persisted. Qualitative evidence points to the psychosocial component as the primary active ingredient, while the economic arm was operationally compromised by the non-delivery of promised capital grants. We interpret the longer-term convergence as durable individual- and interpersonal-level change for treatment women, set against contemporaneous structural interventions in the study region that plausibly affected the control group in particular.

1. Introduction

Intimate partner violence (IPV) is a serious public health problem and human rights issue that affects hundreds of millions of women globally. Worldwide, more than one in four women has experienced some form of physical or sexual IPV in her lifetime (Sardinha et al., 2022; WHO, 2024). Beyond its immediate physical and mental health consequences (Bacchus et al., 2018; Smith et al., 2017), IPV constrains women's ability to take up economic opportunities, participate in civic life, and invest in their children, and imposes substantial costs on survivors, families, and communities (Peterson et al., 2018).¹ Because

IPV is shaped by factors at the individual, interpersonal, community, and societal levels simultaneously, effective prevention is widely understood to require interventions that address multiple drivers at once (Heise, 1998; Ranganathan et al., 2021).

Two broad strands of the empirical literature inform the design of such interventions. The first documents that programs directly targeting women's mental health and relational skills (e.g., cognitive behavioral therapy, group counseling, and couple-based communication training) can reduce IPV in low- and middle-income countries. Ellsberg et al. (2015) synthesize RCT and quasi-experimental evaluations of

[☆] I am extremely grateful to Jonathan Robinson, my advisor, for his tireless support and mentoring over the course of this project. I also sincerely appreciate the guidance and encouragement from Natalia Lazzati, Ajay Shenoy, and Alan Spearot. For helpful discussions and comments, I also thank Shilpa Aggarwal, Dahyeon Jeong, Naresh Kumar, and seminar participants at the KDI School, UCSC Micro Workshop, UC Berkeley Development Lunch, NEUDC 2021, PacDev 2022, MIEDC 2022, and World Bank DIME-KDI School Development Impact Conference 2023. For organizing the data collection, I thank Arja Dayal, Dackermue Dolo, Wilson Dorleleay, Walker Higgins, Andreas Holzinger, Teresa Martens, and Camelia Vasilov at IPA Liberia and Joseph Kamara. I am extremely grateful to all the enumerators who collected this data, though there are too many to list individually. I thank Christopher Johnson, Johnny Jones, Ambullai Perry, and Sayba Tamba at the Liberia National Red Cross Society for their collaboration. Funding for this project was provided by the UK Foreign, Commonwealth & Development Office, awarded through the J-PAL Crime and Violence Initiative, UC Institute on Global Conflict and Cooperation, IPA IPV Initiative, UCSC Economics Department, UCSC Blum Center, and the KDI School of Public Policy and Management. This study has been approved by the UC Santa Cruz Institutional Review Board (Protocol 3307) and the University of Liberia Institutional Review Board (Protocol 19-02-156). Trial is registered at the AEA RCT Registry under AEARCTR-0004488 (Park 2020). Any error is my own.

E-mail address: park@kdis.ac.kr.

¹ According to the U.S. Centers for Disease Control and Prevention, about 35% of female IPV survivors experience some form of physical injury related to IPV (Smith et al., 2017). In our study sample, about 25% of physical/sexual IPV survivors report a physical injury as a direct consequence of a male partner's action.

<https://doi.org/10.1016/j.jdevec.2026.103823>

Received 22 June 2025; Received in revised form 7 May 2026; Accepted 12 May 2026

Available online 21 May 2026

0304-3878/© 2026 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

violence-against-women interventions globally and find that nearly half report significant reductions in physical or sexual IPV. Pooled estimates from Turner et al. (2020) for 33 psychosocial interventions in low- and middle-income countries imply an average IPV-risk reduction of around 27%. The second strand documents that economic interventions alone produce heterogeneous effects: cash transfers and vocational training sometimes reduce IPV by increasing women's outside options and bargaining power (Hidrobo et al., 2016; Haushofer et al., 2019; Roy et al., 2019), but can also trigger male backlash or household conflict in particular subgroups (Angelucci, 2008; Hidrobo and Fernald, 2013).

Combining the two, economic empowerment paired with intensive psychosocial therapy, is theoretically attractive because the psychosocial component provides the interpersonal skills and community-level norm change that economic gains alone may not secure. A small literature evaluates programs that bundle these components. The closest in design is Bandiera et al. (2020), who combined vocational and life-skills training delivered in community group sessions and documented sustained reductions in gender-based violence—but among adolescent girls in Uganda. Other multifaceted programs targeting adults have examined lighter combinations, such as village savings plus short-duration training (Green et al., 2015) or community mobilization (Abramsky et al., 2014). These studies typically observe short- or medium-run outcomes, and there is little evidence on whether benefits from a comparably intensive combined program persist for adult women once the program and its funding end.

This paper evaluates a multifaceted female empowerment program in Monrovia, Liberia, that combines intensive psychosocial therapy with vocational skills training over 12 months. The program (the Women Training and Integration, or WIN, Program), run by the Liberia National Red Cross Society (LNRCS), is considerably more intensive than most interventions in the literature: roughly 1200 contact hours per participant, of which approximately 600–700 are devoted to the psychosocial components (group counseling, cognitive behavioral therapy, relationship management, and stress management) and the remainder to vocational and business-management training. We evaluate the program in a setting where IPV prevalence is among the highest documented worldwide: in our baseline sample, 51% of women report past-year physical or sexual IPV, reflecting Liberia's high overall prevalence and the persistent legacy of two decades of civil conflict (Omanyondo, 2005; Kelly et al., 2021, 2024).

Access to the program was randomized among eligible applicants, stratified by baseline physical/sexual IPV status and civil-war exposure. We observe treatment and control women at baseline (April 2019), at a first endline 12 months after the program ended (April 2021, “Endline 1”), and at a second endline 3.5 years after the program ended (December 2023–January 2024, “Endline 2”).² We pre-specify intimate partner violence and psychological well-being as the two primary outcome domains (Park, 2020), reflecting the ex ante design intent that the intervention affect both through its integrated psychosocial and economic arms. To help interpret the quantitative findings, we also conducted focus group discussions (FGDs) with program participants in February 2023.

We find three main results. First, the program produced large and sustained reductions in IPV for the treatment group. At Endline 1, emotional IPV declined by 17 percentage points (from a control mean of 62%), physical IPV by 19 pp (from 45%), and physical/sexual IPV by 21 pp (from 50%); all three are statistically significant at conventional levels and survive multiple-testing adjustment. These magnitudes correspond to 27–42% relative reductions in IPV prevalence (treatment–control differences expressed as fractions of the control means) and are broadly comparable to the meta-analytic benchmarks discussed above. At Endline 2 the treatment group's IPV levels remain close to

their Endline 1 values, but the treatment–control difference has closed because the control group also experienced substantial reductions. This pattern is crucial for interpretation: the program's gains did not “fade” within treated women, but rather the counterfactual moved. Consistent with a durable treatment-induced shift, perceived community norms against IPV remain significantly lower for treatment women even at 3.5 years (–0.24 SD), while own attitudes converge.

Second, the economic arm of the program, of which promised starter-kit grants were not delivered because of COVID-era funding shortfalls, produced limited and non-durable effects. Short-term expenditures increased by approximately 36% (the only statistically significant short-term economic effect). Effects on labor supply, income, and business ownership were not statistically significant at either endline, and the Endline 1 point estimates are imprecise enough that large effects in either direction cannot be ruled out. Participants reported in the FGDs that the absence of starter capital materially reduced their ability to translate vocational skills into sustained self-employment.

Third, despite the intensity of the psychosocial component, measured effects on standardized symptom-level psychological well-being indices (Hopkins Symptom Checklist-10, Patient Health Questionnaire-9, happiness and life satisfaction questions from the World Values Survey) are null at both endlines. The FGDs nonetheless provide qualitative evidence that participants describe the counseling component as the element of the program they recalled most vividly, and report continuing to apply skills they learned years after the program ended. Together with the program's contact-hour allocation, the FGD evidence supports a behavioral and relational reading of the IPV effects—channels the standardized symptom-level instruments are not designed to detect.

We interpret the longer-term convergence in IPV through an ecological lens that distinguishes individual, interpersonal, community, and societal drivers of violence. At the individual and interpersonal levels, the program's psychosocial component plausibly produced durable improvements in skills and relational dynamics for treatment women, consistent with both the sustained treatment-group IPV reductions and the FGD testimony. At the community and societal levels, several contemporaneous initiatives in Greater Monrovia, including national programs targeting gender norms, cash transfer programs, and pandemic-era gender-based violence (GBV) response programming, likely reached the control group as well.³ National labor-force participation data and IPV prevalence data from the Demographic and Health Surveys (DHS) show limited aggregate change over this period, so we do not attribute the convergence to broad labor-market or price-level shifts. We therefore read the convergence as the product of multiple overlapping community- and societal-level interventions reaching the control group, rather than as a national-level secular trend; we develop this interpretation, and acknowledge its limits, in Section 4.

This study contributes to several strands of literature. First, we add to the literature on IPV and economic empowerment programs (Hidrobo et al., 2016; Bandiera et al., 2020; Haushofer et al., 2019; Roy et al., 2019). A growing body of research suggests that increasing women's financial autonomy through cash transfers or employment programs can reduce IPV, often by shifting household bargaining power (Bobonis et al., 2013; Heath et al., 2020), although the evidence is mixed and effects can run in either direction depending on context (Buller et al., 2018). Roy et al. (2019) provide a particularly relevant point of comparison: in rural Bangladesh they find that cash or food transfers paired with behavior-change communication reduced IPV more than transfers alone, with effects sustained six to ten months post-program. Whether such reductions persist once the financial support ends, and whether the pattern generalizes to longer horizons and

² The original design was to pool three cohorts across 2019–2021. Due to COVID-19 disruptions and related funding constraints, the implementing partner ended up not enrolling new cohorts from 2020. This paper evaluates a single cohort (N=395 at baseline) enrolled in 2019.

³ For example, the World Bank's Liberia Women Empowerment Project (approved in June 2022), which includes an explicit gender-norms component; the Liberia Social Safety Nets Project, which delivered urban cash transfers in Greater Monrovia during this window; and the UN Women and UNFPA GBV response programs that scaled up in response to the documented pandemic-era rise in gender-based violence. Section 4 discusses in more detail.

to interventions with much heavier psychosocial content, remains an open question. Our paper speaks to this debate from an unusual vantage point: the cohort we study received the psychosocial and skills curriculum but not the promised business capital component, and we follow them for 3.5 years post-program. We document large and sustained IPV reductions in the treatment group despite the absence of the economic transfer, suggesting that a substantial share of the program's IPV impact can be delivered through the psychosocial channel alone.

Second, we contribute to a growing literature, mostly in public health, on psychosocial and behavioral interventions for IPV in low- and middle-income countries (Ellsberg et al., 2015; Turner et al., 2020; Ranganathan et al., 2021). Most of the underlying studies follow participants for at most 12 months and few examine economically integrated programs. We extend the evidence base in two directions: by following the same women out to 3.5 years and by embedding the psychosocial component in a substantially longer and more intensive intervention than has typically been evaluated. Closer to our setting in design, Bandiera et al. (2020) document sustained reductions in gender-based violence in BRAC's ELA program among adolescent girls in Uganda, and Green et al. (2015) and Abramsky et al. (2014) evaluate lighter-touch combinations of microenterprise support and community mobilization. Our setting (adult women in established partnerships, with very high baseline IPV exposure) and our intensity (1200 contact hours of which 600–700 are psychosocial) are differentiated from these studies such that the comparison itself is informative.

The paper proceeds as follows. Section 2 describes the setting, intervention, experimental design, and data. Section 3 presents the main results. Section 4 discusses mechanisms. Section 5 concludes.

2. Study design and data

2.1. Setting

This study was conducted in Monrovia, Liberia, where IPV is highly prevalent. According to the 2019–2020 Liberia Demographic and Health Survey (DHS), 35% of ever-partnered women aged 15–49 reported experiencing physical or sexual IPV in the past 12 months. This is significantly higher than the corresponding regional averages for Asian (16%), Latin American (12%), and other African (26%) countries.⁴ The study population targeted by the Red Cross reports even higher levels of IPV. In our baseline survey (April 2019), 51% of respondents reported having experienced physical or sexual IPV in the past year.

There are several possible explanations for the persistently high IPV prevalence in Liberia. One key factor is economic vulnerability, as Liberia remains one of the world's poorest countries.⁵ Another likely factor is historical exposure to violence during Liberia's civil wars (1989–1996 and 1999–2003), which resulted in approximately 250,000 deaths (around 10% of the population at the time) and displaced over one million people. During the war, widespread sexual violence against civilians was used as a weapon of war to terrorize communities (Omalyondo, 2005),⁶ and a WHO report estimates that two in three Liberian women experienced sexual violence during the war. Previous works link wartime sexual violence, forced displacement, and the breakdown of community norms during Liberia's conflict to persistently elevated IPV prevalence decades later (Kelly et al., 2021, 2024), and more broadly find that prolonged exposure to violence can give rise

to a “culture of violence” (Steenkamp, 2005; Moser and Winton, 2002) that reshapes norms around partner violence across generations.

2.2. Intervention

The intervention is a multifaceted female empowerment program called the Women Training and Integration (WIN) Program, which has been administered by the Liberia National Red Cross Society (LNRCS) since 2009. The program targets vulnerable women in informal settlements of Monrovia. To qualify, an applicant must belong to a minimum of three of the eligibility groups in Table A.2. LNRCS staff review the application packets, visit the applicants' communities, and interview neighbors and acquaintances to verify the reported information.

The program's main objective is to improve the participants' livelihoods in multiple dimensions. Specifically, the program aims at the following: (1) To economically empower women so that they can self-sustain themselves and their families; (2) To psychologically empower women so that they can better protect themselves from abuse; (3) To help establish and maintain positive relations with their families and communities; (4) To improve knowledge about and thus access to health care and psychological services.

The program is intensive. It requires a 12-month commitment, with participants present at the program center for 4–5 hours a day (in either a morning or afternoon session), 5 days a week. Total contact time per participant is approximately 1200 hours.

Of these 1200 hours, roughly 600–700 are devoted to the *psychosocial* component: one-to-one and group counseling, thematic group discussions, cognitive behavioral therapy modules, family/couple therapy and mediation, conflict resolution, and stress management. The psychosocial curriculum aims to heal war-related trauma, reduce traumatic stress, mediate family-conflict situations, support coping mechanisms, build self-confidence, and promote peaceful coexistence within families and communities.

The remaining 500–600 hours are split between *vocational-skills training* (approximately 400 hours in one of three tracks: tailoring, baking/catering, or hairdressing/cosmetology) and *business and financial-management modules* (approximately 150 hours, covering client interactions, record-keeping, pricing, saving, and inventory management). Participants choose one vocational track or are assigned one based on capacity constraints. At graduation, beneficiaries are intended to receive business startup kits and small cash grants to help establish their own enterprises; however, due to financial constraints and COVID-related disruptions, LNRCS was not able to deliver these capital grants for the cohort evaluated in this paper.

The program also includes several auxiliary components. Routine health-care check-ups and HIV/AIDS awareness and testing are provided in LNRCS's in-house clinic, and child-care services are available while participants are at the program center. An adult-literacy module trains unschooled participants in basic arithmetic and English reading and writing, with the curriculum aligned to the Ministry of Education's Alternative Learning Curriculum. As a location-based intervention, the program also provides a “safe place” to meet and socialize with other participants, which may itself affect psychological well-being.

As noted above, the planned expansion to additional cohorts was indefinitely suspended due to COVID-19 disruptions and funding shortfalls, and the study sample is limited to the single 2019–2020 cohort. The qualitative evidence presented in Section 4 documents that the non-delivery of the capital grants generated frustration and, in some cases, debt as women borrowed to improvise starter capital.

2.3. Experimental design

The sampling frame consists of women who voluntarily applied to the WIN program and were screened for eligibility by the Red Cross. That is, our sample can be characterized as women who are disadvantaged enough for LNRCS to consider them eligible for the program but at the same time are willing to improve their lives and have high enough agency to apply to such a program.

⁴ Unweighted averages are calculated across countries with available DHS data.

⁵ Liberia's GDP per capita is among the lowest globally (CIAWorldFactbook) with weak institutions, and many lack access to formal education and sustainable economic activities. For example, per one of the UN's Millennium Development Goals, the net primary education enrollment in Liberia was 37% in 2016, while the average of Sub-Saharan African countries was 78% (UNESCOInstituteForStatistics).

⁶ Also see Domingo et al. (2013), Jones et al. (2014), and UN Women (2013).

Several months before program start for every cohort, LNRCS advertises the program in target communities to encourage eligible women to apply. In February 2019, LNRCS received about 600–700 applications in total, and after background checks and verification of the applicants' information, it shared with us a list of 450 eligible applicants divided into the "main" list of 400 and a "backup" list of 50 ranked in the order of eligibility status determined by LNRCS. In conducting the baseline survey, for those we could not reach after numerous attempts, we drew from the backup list in order. In the end, we enrolled 395 respondents for the study and conducted baseline in April 2019,⁷ and randomly assigned 198 to treatment and 197 to control.

Treatment assignment was stratified by two background characteristics collected in the baseline survey: (a) whether having experienced physical or sexual IPV in the past 12 months, and (b) having been affected by the civil war or having family members who have.⁸

Every woman in the treatment group was invited to the program, but some could not be reached or could not participate in the program for other reasons, and 152 women ultimately enrolled. Moreover, due to an administrative error, 2 people from the control group were invited and joined the program.

2.4. Data collection

We collected survey data at three time points: a baseline survey in April 2019 (before randomization and program start), a first endline (Endline 1) in April 2021 (approximately 12 months after the program ended), and a second endline (Endline 2) between December 2023 and January 2024 (approximately 3.5 years after the program ended). All three questionnaires share a common structure covering IPV, psychological well-being, social norms around IPV, labor supply, income, expenditures, transfers, savings, and credit.

Survey mode varied across waves in response to field conditions. The baseline was conducted in person in the study communities in Monrovia. Endline 1, fielded in April 2021, was conducted predominantly in person (348 of 359 tracked respondents), with phone interviews used for the small number of respondents who could not be reached in person (11 of 359). Endline 2 followed the same protocol (252 of 283 respondents interviewed in person; 31 by phone), with phone interviews used primarily for respondents who had relocated outside Greater Monrovia. All regressions control for a binary phone/in-person indicator.

Because the IPV module is sensitive and its validity depends on privacy, we additionally implemented a planned measurement experiment at Endline 1. A random subsample of respondents completed the IPV module via audio computer-assisted self-interviewing (ACASI), while the remaining respondents completed it via face-to-face interviewing (FTFI). Findings from a parallel study (Park et al., 2025) suggest that the two modalities yield systematically different reports in ways that complicate pooling. Accordingly, the main IPV analyses in this paper are restricted to respondents assigned to the FTFI arm at Endline 1. At Endline 2, IPV was measured only via FTFI.

In addition to the surveys, we conducted five sessions of focus group discussions (FGDs) in February 2023 with a randomly selected subset of 20 treatment-group participants. The FGDs were moderated by trained female facilitators, covered themes of program experience, perceived program impacts on relationships and livelihoods, and experiences since program completion, and were audio-recorded and transcribed for analysis. Throughout the paper we use the FGDs qualitatively, to help interpret the quantitative findings.

⁷ We had completed full interviews with 400 women, but LNRCS later decided to drop anyone under 17 from the sample due to potential conflict with school enrollment.

⁸ Instances include: relocation, becoming disabled/amputated, family members being killed/dead.

2.5. IPV measurement and safety protocols

We measure IPV outcomes using the WHO's Violence Against Women module,⁹ which has become a standard instrument for IPV measurement across diverse contexts. The module consists of a series of binary questions, each describing a specific IPV-related episode, providing respondents with multiple opportunities to report violence. We group the responses into four categories: controlling behavior, emotional IPV, physical IPV, and sexual IPV. The recall period is the past 12 months prior to the survey date. Appendix B provides a more comprehensive description of the questionnaire.

We followed the WHO's ethics protocol for IPV research (WHO, 2016). The study was approved by the institutional review boards of the University of California, Santa Cruz, and the University of Liberia. We hired only female enumerators and trained them both to conduct the interviews safely and to manage the emotional demands of the work. The full survey was conducted privately, with no third party present beyond the enumerator and the respondent; in the IPV module specifically, enumerators were trained to switch to non-sensitive subjects if the interview was interrupted or eavesdropped. Although respondents had completed an informed-consent procedure at the start of the survey that already covered the IPV module, we reiterated consent immediately before the IPV section. Finally, every respondent who went through the IPV questionnaire received an information sheet listing services available to women experiencing IPV, including contact details for organizations that provide help, regardless of whether she reported any IPV experience.

2.6. Baseline summary statistics

Table 1 presents baseline summary statistics. The average age of women in the control group is about 29 years. They completed 7 years of education, on average, and about two-thirds of our sample have completed primary school, while only 25% of women have completed secondary school.

For the IPV questions, we restrict the sample to those who are currently partnered or have had an intimate partner 12 months prior to the survey, and the mean for this indicator at baseline was 92%.¹⁰

In Panel B we find that our sample had minimal access to their own income source or labor force participation opportunities. Only 11% report having any job, and 25% are self-employed. The average income is a mere 8 dollars per month, with many zeros in the extensive margin. The mean for spouse's income is twice as large (\$19). While our measures of income might not be exhaustive themselves, the mean differences suggest that the women in our sample were not financially independent at baseline.

The baseline prevalence of IPV is very high. About 59% of women reported having experienced emotional IPV, while the figure for the more severe form of IPV (physical or sexual) is slightly smaller (51%). This rate is much higher than the national average reported in the Liberia DHS surveys, where the corresponding figures are 35% and 35% respectively in the 2019–2020 report. There could be two possible explanations. One is that our sample was selected by Red Cross in a way to be characterized as vulnerable, and one eligibility criterion was having experienced domestic abuse (Table A.2). Another is the difference in survey tool between our baseline and the Liberia DHS 2019–2020. While our study uses the identical questionnaire to the DHS's Domestic Violence Module, at our baseline IPV was measured solely in audio computer-assisted self-interviewing (ACASI), and DHS data are measured via traditional face-to-face interviewing (FTFI). In light of the findings in our sister project in rural Liberia and Malawi (Park et al., 2025), the reported differences could be due to differing measurement modality, either through enhanced confidentiality or increased

⁹ https://www.who.int/gender/violence/who_multicountry_study/Annex3-Annex4.pdf.

¹⁰ We later show in Table A.3 that this indicator is statistically balanced between treatment and control at both endlines.

Table 1
Baseline summary statistics and randomization check.

	(1) Control Mean [SD]	(2) Treatment - Control
Panel A: Demographics		
Age	28.98 [7.29]	1.36* (0.73)
Years of education	7.27 [4.11]	0.45 (0.40)
= 1 if completed primary school	0.66	0.06 (0.05)
= 1 if completed secondary school	0.25	0.01 (0.04)
= 1 if currently partnered or had partner past year	0.92	-0.00 (0.03)
Panel B: Self income and labor supply		
= 1 if has own income source (past month)	0.34	0.06 (0.05)
= 1 if operated own business (past month)	0.25	0.04 (0.04)
= 1 if had any other temporary/permanent job (past month)	0.11	0.01 (0.03)
Total income (past month, USD)	8.38 [27.57]	3.36 (3.09)
Panel C: Household economic well-being		
Spouse's income (past month, USD)	19.06 [39.56]	2.11 (4.05)
Per capita expenditure (past month, USD)	26.76 [25.54]	1.65 (2.63)
Net value of physical assets (USD)	316.32 [1,282.83]	80.88 (133.55)
Panel D: Intimate partner violence		
= 1 if experienced the following (past 12 months):		
Controlling behavior	0.83	0.03 (0.04)
Emotional IPV	0.59	0.00 (0.05)
Physical IPV	0.50	-0.01 (0.05)
Sexual IPV	0.16	0.03 (0.04)
Physical or sexual IPV	0.51	-0.01 (0.05)
Emotional, physical, or sexual IPV	0.67	-0.02 (0.05)

Note: Baseline survey conducted in April 2019. N = 395 women, with 198 randomly assigned to treatment and 197 to control, stratified by baseline physical/sexual IPV experience and civil-war exposure. Column (1) reports control-group means with standard deviations in brackets; column (2) reports the treatment-control difference with heteroskedasticity-robust standard errors in parentheses. The IPV module at baseline was administered via audio computer-assisted self-interviewing (ACASI) and is restricted to women currently partnered or partnered in the past 12 months. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

measurement error. Yet, the control group's IPV rates at our first endline measured in FTFI only are still high—62% for emotional IPV, 45% for physical IPV, and 24% for sexual IPV (Table 2).

2.7. Attrition balance

In Table A.3, we check balance for two compliance measures: columns (1) and (3) show whether we were able to reach the respondent and complete the follow-up survey itself, and columns (2) and (4) refer to whether she was eligible for the IPV section at endline. Given our IPV questions have a recall period of 12 months, we administered the IPV module only to those who are currently partnered or have been so in the past 12 months. Since the IPV analysis is indeed constrained to only those who went through the IPV questionnaire at all, it is necessary to check for any differential attrition in partner status. In addition, given that often in developed countries, IPV survivors are encouraged to leave the violent partner, this is also a meaningful outcome that would show how women in our study select in or out of a relationship at all.

Overall, at the first endline (about 2 years since baseline), we successfully tracked about 91% of baseline respondents, and find no

significant differences between treatment and control either in survey completion or IPV module eligibility (columns 1 and 2). By the second endline (about 4.5 years since baseline), attrition increased to roughly 30% overall. Survey completion remains statistically balanced between treatment and control (column 3), and IPV-module eligibility differs by only about one percentage point (column 4), again statistically insignificant.

For analysis, the estimation samples are: Baseline N=395; Endline 1 full survey N=359, of whom 169 are currently or recently partnered *and* in the FTFI arm of the measurement experiment (the IPV estimation sample); Endline 2 full survey N=283, of whom 245 are the IPV estimation sample (all FTFI at Endline 2). These counts are reproduced in the notes to each affected table.

3. Results

3.1. Program effects on IPV

Our preferred specification is an ANCOVA model:

$$Y_i = \alpha + \beta W I N_i + \gamma Y_{0i} + \delta phone_i + \phi_s + \varepsilon_i, \quad (1)$$

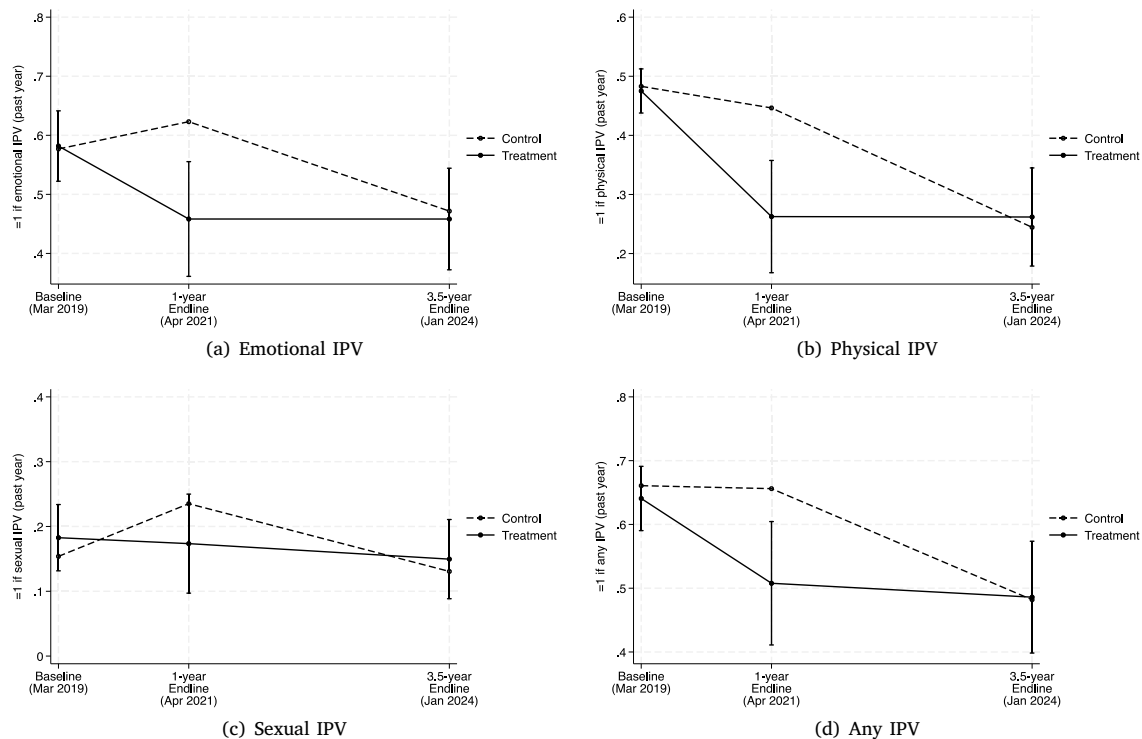


Fig. 1. Intimate partner violence (IPV) indices across time.

Note: Each panel plots mean rates of the indicated IPV indicator (Panel a: emotional IPV; Panel b: physical IPV; Panel c: sexual IPV; Panel d: any IPV) for the treatment and control groups at baseline (April 2019), Endline 1 (April 2021), and Endline 2 (December 2023–January 2024). Vertical bars denote 95% confidence intervals based on heteroskedasticity-robust standard errors; underlying regressions include strata fixed effects. The treatment intervention ran from April 2019 through March 2020. The IPV module was administered via audio computer-assisted self-interviewing (ACASI) at baseline and via face-to-face interviewing (FTFI) at both endlines; the random Endline-1 ACASI subsample is excluded from these plots. The IPV sample restricts to women currently partnered or partnered in the past 12 months. Sample sizes: N=395 at baseline; N=169 at Endline 1; N=245 at Endline 2.

where Y_i is the outcome of interest for individual i , WIN_i is the treatment-assignment indicator, Y_{0i} is the baseline measurement of the outcome (with a missing-indicator variable for respondents without baseline values), $phone_i$ is an indicator for phone-based interviewing, and ϕ_s denotes strata fixed effects. Standard errors are heteroskedasticity-robust. For IPV outcomes that were not collected at baseline (e.g., the physical/sexual composite and the IPV-frequency sum), we control instead for baseline measurements of the component variables.

To address multiple-hypothesis testing across the IPV outcomes within each wave, we report sharpened two-stage q -values following Benjamini et al. (2006) as applied by Anderson (2008). The family of tests in each panel comprises the six IPV outcomes shown in the columns of Table 2 (controlling behavior, emotional, physical, sexual, physical-or-sexual composite, and any IPV); the procedure controls the false discovery rate while exploiting the data-driven estimate of the number of true nulls. Sharpened q -values are reported as a separate row in the IPV regression tables.

Fig. 1 plots the time path of the four primary IPV indicators (emotional, physical, sexual, and any IPV) for treatment and control at baseline (April 2019, ACASI), Endline 1 (April 2021, FTFI sub-sample), and Endline 2 (December 2023–January 2024, FTFI). Two features are visible. First, between Endline 1 and Endline 2 the treatment group's IPV rates remain close to their Endline 1 values; emotional IPV in particular is essentially flat in the treatment group. Second, the control group's IPV rates fall substantially between the two endlines, from 62% to approximately 46% for emotional IPV, and from 45% to approximately 23% for physical IPV. The Endline 1 treatment–control gap is closed at Endline 2 not because the program's effects dissipated in treated women but because the control group caught up.

Table 2 presents the corresponding regression estimates. At Endline 1 (Panel A), emotional IPV declines by 17 percentage points ($p < 0.05$), physical IPV by 19 percentage points ($p < 0.01$), physical/sexual IPV by 21 percentage points ($p < 0.01$), and any IPV by 14 percentage points ($p < 0.05$); the estimate on sexual IPV (-0.07) is not statistically distinguishable from zero. All four significant effects survive multiple-testing adjustment via sharpened q -values at the $q < 0.05$ level. Relative to control means of 45%–66%, these magnitudes correspond to approximately 21–42% relative reductions in IPV prevalence (treatment–control difference divided by the control mean). They sit at the upper end of the range reported in the public-health evaluation literature on psychosocial and combined IPV interventions in low- and middle-income countries (e.g., Ellsberg et al. (2015), Turner et al. (2020)), in which the average pooled effect is around 27% across heterogeneous trials. Results using frequency-weighted IPV indices are in the same direction (Table A.4). Table A.6 and Fig. A.2 report individual-question binary indicators, and the pattern confirms that the index-level reductions are driven by physical/sexual items.

At Endline 2 (Panel B), the regression-adjusted treatment–control differences are small and not statistically significant. Point estimates on emotional, physical, and any IPV are close to zero. Section 4 discusses the mechanisms we believe account for this convergence.

The pre-registered stratification variable was baseline physical/sexual IPV status. Table A.7 reports heterogeneous treatment effects on IPV by this dimension. At Endline 1, treatment effects are attenuated (though not reversed) among women with highest baseline IPV (interaction coefficients of roughly $+0.27$ – 0.28 , significant for emotional, sexual, and any IPV); at Endline 2 the interactions are close to zero. One reading of the attenuation pattern at Endline 1 is that the most severely affected couples, who likely face the deepest constraints on relational

Table 2
Program effects on IPV indices.

	(1) Controlling Behavior	(2) Emotional Violence	(3) Physical Violence	(4) Sexual Violence	(5) Physical or Sexual	(6) Any IPV
Panel A. 1st Follow-up (1 year post program end)						
WIN treatment	−0.02 (0.06)	−0.17** (0.07)	−0.19*** (0.07)	−0.07 (0.06)	−0.21*** (0.07)	−0.14** (0.07)
Sharpened <i>q</i> -value	0.236	0.014	0.008	0.099	0.008	0.021
Control mean	0.80	0.62	0.45	0.24	0.50	0.66
Observations	169	169	169	169	169	169
Panel B. 2nd Follow-up (3.5 years post program end)						
WIN treatment	−0.06 (0.05)	−0.01 (0.06)	0.01 (0.06)	0.01 (0.04)	0.01 (0.06)	0.01 (0.06)
Sharpened <i>q</i> -value	0.895	0.895	0.895	0.895	0.895	0.895
Control mean	0.80	0.46	0.23	0.12	0.26	0.47
Observations	245	245	245	245	245	245

Note: ANCOVA estimates from Eq. (1). Panel A reports Endline 1 results (April 2021, $N = 169$ women in the FTFI arm of the measurement experiment); Panel B reports Endline 2 results (December 2023–January 2024, $N = 245$, all FTFI). The IPV sample is restricted to women currently partnered or partnered in the past 12 months. Recall period is the 12 months prior to each survey. Regressions include baseline measurement of the outcome (or, for composites not measured at baseline, the corresponding component variables), strata fixed effects, and a control for whether the survey was conducted by phone. Heteroskedasticity-robust standard errors in parentheses. The row labeled “*q*-value” reports sharpened two-stage *q*-values (Benjamini et al., 2006, as applied by Anderson, 2008) for the family of six IPV outcomes within each panel. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

change, may benefit less from the program’s standard psychosocial dosage and could need a different or more intensive intervention; given limited power, we do not pursue this interpretation further. Heterogeneity by baseline civil-war exposure (Table A.8) is uniformly small and statistically insignificant. The heterogeneity patterns are consistent with the program being effective across the full range of baseline IPV severity and war exposure in our sample, rather than being driven by a narrow subgroup.

3.2. Program effects on psychological well-being

Given the substantial psychosocial component of the program, we separately examine its effects on standardized measures of psychological well-being. We constructed three indices: (i) a distress index based on the ten-item Hopkins Symptom Checklist (HSCL-10), which is widely used in clinical and epidemiological settings to measure symptoms of anxiety and depression; (ii) a depression index based on the nine-item Patient Health Questionnaire (PHQ-9); and (iii) a happiness index constructed from the Happiness and Well-being module of the World Values Survey, standardized following Anderson (2008). We also report an overall psychological well-being composite that pools the three.

Table 3 presents the estimates. Point estimates move in the expected direction at Endline 1 (lower distress, higher happiness), but none is statistically distinguishable from zero. At Endline 2, 3.5 years after program completion, the estimates remain close to zero and statistically insignificant for each of the two measures and for the composite index. We read the null on standardized symptom-level instruments together with the sustained IPV reductions in Section 3.1 as consistent with a behavioral- and relational-channel interpretation of the program (Section 4) — effects that the HSCL-10, PHQ-9, and WVS happiness instruments are not designed to detect.

This reading is supported by a growing literature on psychotherapy interventions in low- and middle-income countries showing that symptom-level improvements in standardized instruments can fade within a year of the last therapy contact even when behavioral effects persist (Blattman et al., 2017; Haushofer et al., 2020). Because our first post-program measurement was taken twelve months after program completion, we cannot rule out short-lived symptom effects during or immediately after the intervention.

3.3. Social norms around IPV

A central mechanism through which the program could affect IPV is a shift in normative beliefs about the acceptability of violence.

Public-health scholarship identifies such norms as a key target for behavioral interventions (Ranganathan et al., 2021). We measured justifiability of IPV through seven questions asking whether a husband is justified in hitting or beating his wife in specified circumstances (e.g., if she argues with him, refuses sex, or neglects the children). The same items were asked both for the respondent’s own beliefs and for what she believed “most people in her community” think. We aggregated responses into composite indices using inverse-covariance weighting.

Table 4 reports the composite indices and the self-referenced items; Table 5 reports the community-referenced items. At Endline 1, both the self-referenced and the community-referenced composite indices move downward by roughly 0.20 standard deviations. The two effect sizes are statistically indistinguishable from each other, and we do not read the small numerical gap between them as substantively meaningful. The reductions load on the items with the highest baseline acceptance (“neglecting children”, “arguing with the husband”, and “going out without informing the husband”) on both the self-referenced side (Table 4) and the community-referenced side (Table 5), rather than on items already near zero at baseline. The magnitude is consistent with the program’s explicit curriculum content on IPV acceptability, self-worth, and healthy relationships.

The more informative finding emerges at Endline 2. The self-referenced composite index has reverted close to the control mean, but the community-referenced composite remains approximately 0.24 standard deviations lower in the treatment group and is statistically significant (Table 4, Table 5). The item-level decomposition in Table 5 confirms that this longer-term effect loads on items similar to those in the short-run effect, which rules out the possibility that the persistence is being driven by a few outlier items or by reshuffling across items. In other words, 3.5 years after program completion, treatment women’s own stated attitudes had converged to control levels, while their perception of what is considered acceptable in their community had not, and the gap was driven by the same dimensions of acceptability that moved at Endline 1.

One interpretation, supported qualitatively in the FGDs (Section 4), is that the program’s community-level rationale (the curriculum spent substantial time on how IPV is normalized and how participants can push back in their social networks) continued to inform how treatment women read the normative environment around them even after individual attitudes converged. An alternative interpretation is that control women also shifted their own views over the 3.5-year window in response to overlapping community-level interventions, while the treatment group’s perception of community norms, anchored by the

Table 3
Program effects on psychological well-being.

	(1) HSCL-10 (mean)	(2) PHQ-9 (sum)	(3) WVS (z-score)	(4) Composite (z-score)
Panel A. 1st Follow-up (1 year post program end)				
WIN treatment	−0.01 (0.05)		0.05 (0.10)	0.05 (0.11)
Sharpened <i>q</i> -value	0.853		0.853	0.853
Control mean	2.01		0.00	−0.00
Observations	359		359	359
Panel B. 2nd Follow-up (3.5 years post program end)				
WIN treatment		−0.22 (0.59)	0.01 (0.11)	0.02 (0.11)
Sharpened <i>q</i> -value		0.960	0.960	0.960
Control mean		16.87	−0.00	−0.00
Observations		283	283	283

Note: ANCOVA estimates from Eq. (1). Panel A reports Endline 1 results (April 2021, $N = 359$); Panel B reports Endline 2 results (December 2023–January 2024, $N = 283$). HSCL-10 is the ten-item Hopkins Symptom Checklist; PHQ-9 is the nine-item Patient Health Questionnaire; the WVS happiness measure is standardized following Anderson (2008). PHQ-9 was not administered at the 1st Follow-up and HSCL-10 was not administered at the 2nd Follow-up. The composite index pools the two available measures in each wave using inverse-covariance weights. Regressions include baseline measurement of the outcome (where available), strata fixed effects, and a control for whether the survey was conducted by phone. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 4
Program effects on justifiability of IPV: self-referenced items.

	(1) = 1 if husband is justified to beat/hit wife when she: Argues w/ husband	(2) Goes out w/o telling	(3) Does not care children	(4) Burns food	(5) Financial pressure	(6) Refuses sex	(7) = 1 if husband is justified to force sex	(8) Composite (z-score)
Panel A. 1st Follow-up (1 year post program end)								
WIN treatment	−0.05* (0.03)	−0.03 (0.02)	−0.07** (0.03)	−0.01 (0.02)	−0.02 (0.02)	−0.01 (0.02)	0.01 (0.02)	−0.20** (0.09)
Sharpened <i>q</i> -value	0.182	0.232	0.088	0.680	0.232	0.680	0.680	0.117
Control mean	0.08	0.07	0.12	0.03	0.03	0.02	0.02	−0.03
Observations	359	359	359	359	359	359	359	359
Panel B. 2nd Follow-up (3.5 years post program end)								
WIN treatment	0.04 (0.04)	0.02 (0.04)	−0.01 (0.04)	0.05* (0.03)	0.06** (0.03)	0.02 (0.03)	0.02 (0.02)	0.04 (0.13)
Sharpened <i>q</i> -value	0.642	0.740	0.835	0.397	0.380	0.740	0.740	0.835
Control mean	0.09	0.09	0.12	0.04	0.04	0.04	0.04	0.02
Observations	283	283	283	283	283	283	283	283

Note: Panel A reports Endline 1 results (April 2021, $N = 359$); Panel B reports Endline 2 results (December 2023–January 2024, $N = 283$). The last column reports an inverse-covariance weighted composite of the seven items; the remaining columns report each item separately. Items are binary indicators for whether the respondent considers physical violence justified in the specified circumstance. Community-referenced versions of the same items are reported in Table 5. Regressions include baseline measurement of the outcome, strata fixed effects, and a control for whether the survey was conducted by phone. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 5
Program effects on justifiability of IPV: community-referenced items.

	(1) = 1 if husband is justified to beat/hit wife when she: Argues w/ husband	(2) Goes out w/o telling	(3) Does not care children	(4) Burns food	(5) Financial pressure	(6) Refuses sex	(7) = 1 if husband is justified to force sex	(8) Composite (z-score)
Panel A. 1st Follow-up (1 year post program end)								
WIN treatment	−0.04 (0.05)	−0.06 (0.05)	−0.11** (0.04)	−0.09** (0.03)	−0.02 (0.03)	−0.07* (0.04)	−0.06** (0.03)	−0.21** (0.09)
Sharpened <i>q</i> -value	0.357	0.178	0.035	0.035	0.361	0.076	0.069	0.055
Control mean	0.30	0.30	0.27	0.17	0.13	0.16	0.14	−0.02
Observations	359	359	359	359	359	359	359	359
Panel B. 2nd Follow-up (3.5 years post program end)								
WIN treatment	−0.10 (0.06)	−0.03 (0.06)	−0.09 (0.05)	−0.07* (0.04)	−0.07 (0.05)	−0.07 (0.05)	−0.08* (0.05)	−0.24** (0.11)
Sharpened <i>q</i> -value	0.150	0.557	0.150	0.150	0.150	0.191	0.150	0.150
Control mean	0.47	0.40	0.33	0.20	0.22	0.25	0.22	−0.03
Observations	283	283	283	283	283	283	283	283

Note: Panel A reports Endline 1 results (April 2021, $N = 359$); Panel B reports Endline 2 results (December 2023–January 2024, $N = 283$). Items are binary indicators for whether the respondent believes most people in her community consider physical violence justified in the specified circumstance. The last column is an inverse-covariance weighted composite of the seven items. Specification matches Table 4. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

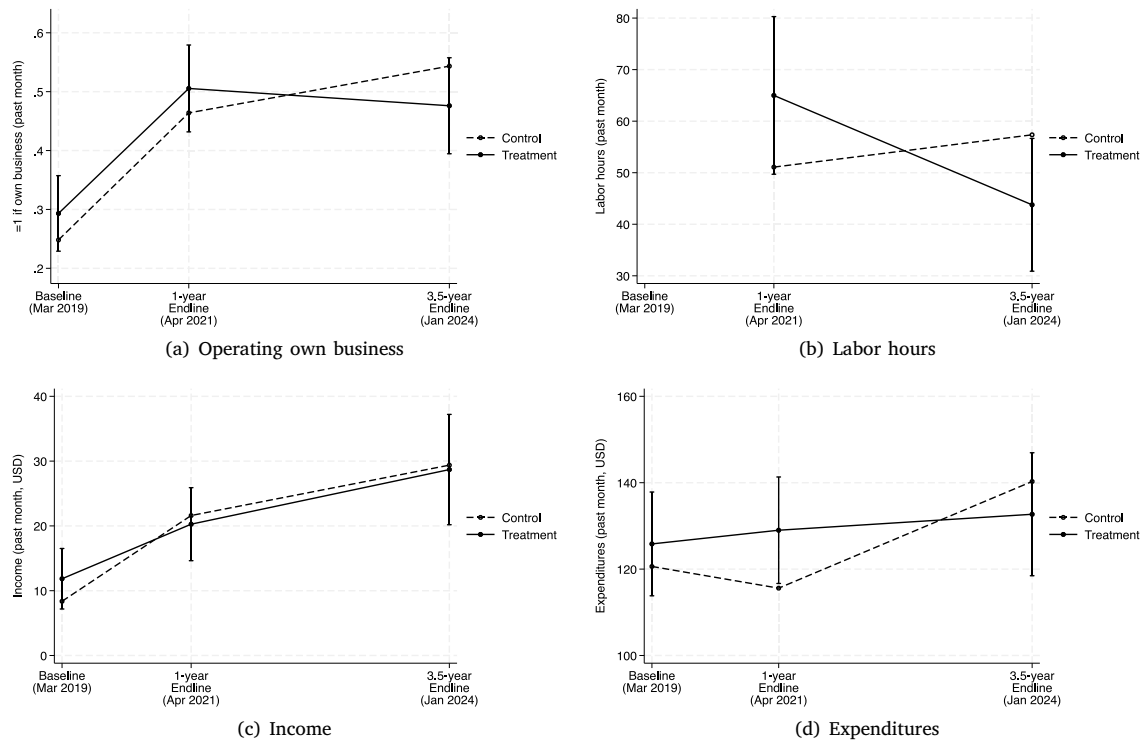


Fig. 2. Labor supply and economic welfare across time.

Note: Each panel plots mean values of the indicated economic outcome (Panel a: indicator for operating own business; Panel b: labor hours in the past month; Panel c: income in the past month, USD; Panel d: expenditures in the past month, USD) for the treatment and control groups at baseline (April 2019), Endline 1 (April 2021), and Endline 2 (December 2023–January 2024). Vertical bars denote 95% confidence intervals based on heteroskedasticity-robust standard errors; underlying regressions include strata fixed effects. The treatment intervention ran from April 2019 through March 2020. Sample sizes: $N=395$ at baseline; $N=359$ at Endline 1; $N=283$ at Endline 2.

program content, remained more pessimistic. We cannot fully separate these two mechanisms.

Regardless of interpretation, the persistent gap in perceived community norms is the cleanest longer-term effect we detect and is consistent with the pattern of sustained IPV reductions in treatment women discussed in Section 3.1. We note that “community norms” as measured here is each respondent’s self-reported perception, not an independently measured community-level outcome; we therefore read the persistent treatment effect on this measure as a durable shift in treatment women’s beliefs about the normative environment rather than as direct evidence of a change in that environment itself.

3.4. Program effects on labor supply, business ownership, and expenditures

We estimate the program’s effects on labor supply, self-employment, income, and expenditures using the same ANCOVA specification as Eq. (1). Results are reported in Table 6 (labor supply and business ownership) and Table 7 (income and expenditures). Time paths for the primary economic outcomes are shown in Fig. 2.

At Endline 1, the only statistically significant economic effect is on household expenditures, which rise by approximately 36% relative to the control mean. Point estimates on self-employment participation, hours of self-employment, and total labor hours are in the direction of a modest increase in the treatment group but are not statistically distinguishable from zero. Standard errors on total labor hours (in the range of seven to ten hours per month) imply that, with our Endline 1 sample of 359 respondents, the minimum detectable effect on labor supply is roughly 14–20 h per month; smaller effects cannot be ruled out. Winsorizing labor hours at the 5th and 95th percentiles (Table A.9) does not meaningfully change the point estimates.

At Endline 2, none of the labor-supply or business-ownership estimates is statistically significant. The point estimate on expenditures reverses sign and is significant on conventional standard errors, though it does not survive multiple-testing adjustment.

The economic results are therefore narrow: a short-run expenditure increase that is not matched by a detectable increase in earnings or labor supply, and no statistically significant improvements on any economic outcome at 3.5 years. We interpret these estimates in light of an operational compromise in the economic arm of the program: the starter-kit grants promised to accompany the WIN program were not delivered for the cohort we study, and FGD participants consistently described this as a material barrier to translating training into sustained self-employment. We discuss the qualitative and contextual interpretation of these patterns in Section 4.

4. Discussion of mechanisms

Two features of our results call for interpretation. First, the treatment group’s IPV rates remain close to their Endline 1 values at 3.5 years, yet the Endline 2 treatment–control difference has closed because the control group also experienced substantial reductions. Second, the one unambiguous longer-term gap is on *perceived community norms* rather than on individual attitudes or economic outcomes. We organize the discussion around Heise’s (1998) ecological framework, which distinguishes individual, interpersonal, community, and societal-level drivers of IPV, and we combine this with qualitative evidence from our 2023 focus group discussions and with contemporaneous data on the policy environment in Liberia.

4.1. Qualitative evidence from focus group discussions

We conducted five focus group discussions (FGDs) in February 2023 with a randomly selected subset of 20 treatment-group participants, moderated by trained female facilitators. The FGDs were not designed to identify causal effects; we draw on them to interpret the quantitative patterns above.

Three themes are consistent across the five groups. First, participants describe the psychosocial component as the element of the

Table 6

Program effects on labor supply and business ownership.

	(1) Self employment = 1 if any	(2) hours	(3) Casual labor = 1 if any	(4) hours	(5) Other income = 1 if any	(6) hours	(7) Total = 1 if any	(8) hours
Panel A. 1st Follow-up (1 year post program end)								
WIN treatment	0.02 (0.05)	15.70 (9.71)	-0.03 (0.03)	1.36 (1.69)	-0.05 (0.03)	-3.65 (4.69)	-0.05 (0.05)	13.18 (10.22)
Sharpened <i>q</i> -value	0.651	0.499	0.499	0.499	0.499	0.499	0.499	0.499
Control mean	0.46	38.38	0.08	1.34	0.12	11.36	0.63	51.08
Observations	359	359	359	359	359	359	359	359
Panel B. 2nd Follow-up (3.5 years post program end)								
WIN treatment	-0.08 (0.06)	-7.52 (9.80)	0.04 (0.04)	0.87 (0.80)	-0.05 (0.04)	-7.92* (4.21)	-0.05 (0.06)	-15.28 (10.27)
Sharpened <i>q</i> -value	0.417	0.444	0.417	0.417	0.417	0.417	0.444	0.417
Control mean	0.54	44.66	0.09	1.20	0.14	11.55	0.67	57.40
Observations	283	283	283	283	283	283	283	283

Note: ANCOVA estimates from Eq. (1). Panel A reports Endline 1 results (April 2021, N = 359); Panel B reports Endline 2 results (December 2023–January 2024, N = 283). Recall period for labor outcomes is the past month. Regressions include baseline measurement of the outcome, strata fixed effects, and a control for whether the survey was conducted by phone. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 7

Program effects on income and expenditures.

	(1) Expenditure	(2) Income	(3) Net Wealth
Panel A. 1st Follow-up (1 year post program end)			
WIN treatment	8.90*** (2.69)	-2.80 (4.07)	58.43 (100.41)
Sharpened <i>q</i> -value	0.002	0.374	0.374
Control mean	24.81	21.71	453.37
Observations	359	359	359
Panel B. 2nd Follow-up (3.5 years post program end)			
WIN treatment	-7.70** (3.83)	-0.54 (6.11)	165.91 (165.02)
Sharpened <i>q</i> -value	0.137	0.929	0.473
Control mean	36.37	29.65	580.75
Observations	283	283	283

Note: ANCOVA estimates from Eq. (1). Panel A reports Endline 1 results (April 2021, N = 359); Panel B reports Endline 2 results (December 2023–January 2024, N = 283). Income and expenditure outcomes refer to the past month and are reported in USD; expenditures are per capita. Regressions include baseline measurement of the outcome, strata fixed effects, and a control for whether the survey was conducted by phone. Heteroskedasticity-robust standard errors in parentheses. Fig. A.1 plots the full expenditure distributions by treatment status and wave. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

program that they recall most vividly. They recount skills acquired during the one-to-one counseling, group therapy, and conflict-management modules (anger control, de-escalation of confrontations, non-provocative expression of needs, self-worth, and partner communication), and report continuing to apply them years after the program ended. Many explicitly attribute reductions in the frequency and severity of their disputes with partners to these skills. We treat this testimony as the standard mixed-methods complement to the quantitative results: it is consistent with the sustained treatment-group IPV effects in Section 3.1 and suggests that any program-induced behavioral changes are likely relational rather than symptomatic, which in turn makes sense of the null on the standardized psychological well-being indices in Section 3.2.

Second, participants describe the cancellation of the promised starter-kit grants as the central operational disappointment. Several women report having borrowed money from neighbors, savings groups, or informal lenders to purchase the tools or materials needed to start a small business based on the vocational training they had received, and describe accumulated debt and shuttered businesses as a consequence. This qualitative account is consistent with the quantitative pattern on economic outcomes (Section 3.4): a short-run bump in expenditures not matched by a detectable increase in earnings or labor supply, fading at 3.5 years. It is also consistent with the broader literature showing that vocational and microenterprise training in low-income settings

typically does not translate into sustained self-employment without an accompanying capital or asset component (Blattman et al., 2014; Blattman and Ralston, 2015; Blattman et al., 2016; Attanasio et al., 2017; Bandiera et al., 2017; McKenzie and Puerto, 2021; Alfonsi et al., 2020; Bandiera et al., 2020).

Third, participants describe a change in their role within their communities. They report becoming more vocal about IPV within their families and neighborhoods, intervening in disputes, advising younger women, and naming violence as violence in conversations with peers. Several participants also remark that, while their own attitudes have shifted, they perceive community attitudes as having shifted less, in the sense that what they themselves regard as acceptable diverges from what they believe their neighbors regard as acceptable. This qualitative observation is the FGD counterpart of the quantitative result in Section 3.3: a persistent longer-term gap between treatment and control in perceived community norms, even as self-reported attitudes converge.

4.2. Interpretation within an ecological framework

Heise's ecological model separates four concentric levels at which IPV risk is determined: the individual (personal history, mental health, substance use), the interpersonal (couple-level dynamics and household bargaining), the community (neighborhood networks, institutions, local norms), and the societal (laws, national policies, cultural norms). We use this structure to organize how the WIN program plausibly operated at each level and where the Endline 2 convergence most likely originates.

At the *individual* and *interpersonal* levels, the program appears to have produced durable change for treatment women. The psychosocial curriculum targeted these layers directly (through therapy, skill-building, and couple-relationship content), and participants report in the FGDs that these skills continued to shape how they managed disputes years after the program ended. The sustained treatment-group IPV rates between Endline 1 and Endline 2 are consistent with this interpretation: treated women are not converging downward; their measured IPV levels are approximately flat.

At the *community* and *societal* levels, several forces active during the 2019–2024 study period plausibly affected the control group as well, providing a candidate mechanism for the control-side reductions in IPV. The World Bank's Liberia Women Empowerment Project (P175934), approved in June 2022, explicitly includes a gender-norms and GBV-prevention component operating in urban Liberia. The Liberia Social Safety Nets Project (P175502) began disbursing urban cash transfers to poor households in Greater Monrovia during the same window, and the international literature documents that cash-transfer programs often reduce IPV as a side effect (Hidrobo et al., 2016; Haushofer et al.,

2019). UN Women and UNFPA scaled up GBV-response programming in Liberia during the COVID-19 period in response to the documented pandemic-era rise in gender-based violence. It is not possible in our data to attribute control-side reductions to any single one of these initiatives, but their joint coverage of the study communities during the 2021–2023 window is a plausible reason the control group's IPV levels declined.

Aggregate indicators are consistent with this reading. National labor-force participation, as reported in the World Development Indicators, changed little between 2021 and 2023 in Liberia, and aggregate IPV prevalence in the 2019–2020 Liberia DHS (the most recent available at the time of writing) was close to the 2007 DHS estimate. We therefore do not attribute the convergence to broad labor-market shifts or to a large secular decline in IPV at the national level. The more likely explanation is the concentrated coverage of community- and societal-level programming in the specific communities from which both treatment and control women were drawn.

The one outcome that is *not* convergent at 3.5 years is perceived community norms. The program devoted curriculum time explicitly to how IPV is normalized, how communities reinforce it, and how participants can challenge these norms in their social networks. FGD testimony suggests that treatment women continued to read their community through this more critical lens even as their own attitudes on direct-justification items converged to control levels. We interpret the persistent gap in perceived community norms at Endline 2 (Section 3.3) as the clearest longer-term evidence we have that the program shifted an interpretive frame at the community level in treated women, separately from any shift in individual attitudes.

4.3. Implications and scope

Three implications follow from this interpretation. First, the fact that treatment–control IPV differences close at 3.5 years is *not* evidence that the program's effects faded: the treatment group is approximately flat and the control group moves toward it. This distinction is difficult to detect without two post-program survey waves and without contextual data on the policy environment, and it has direct implications for how the literature reads longer-term follow-up evaluations of IPV interventions. Second, the program produced sustained IPV reductions in treated women at magnitudes broadly comparable to the meta-analytic benchmark for psychosocial interventions, even though the capital-grants component was not delivered. Together with the contact-hour allocation, the FGD testimony, and the null on the standardized symptom-level well-being indices, this pattern points to the IPV effect being carried by the psychosocial channel. What our design can say is that adding the canceled capital component would not have been necessary to obtain IPV reductions of this size; what it cannot say is the marginal value of either component on its own. Third, the persistent effect on perceived community norms, measured 3.5 years after program completion, is consistent with a reading in which the program's durable reach is interpretive rather than symptomatic: treated women do not report lower measured distress or higher measured happiness, but they do report a different reading of what their community finds acceptable.

The principal limit of this interpretation is that we cannot directly attribute the control-side reduction in IPV to any specific contemporaneous program; the evidence is circumstantial, drawing on coincident dates of known initiatives and on the absence of aggregate labor-market or price-level changes that would otherwise explain the pattern.

5. Conclusion

We evaluated a multifaceted female empowerment program in Monrovia, Liberia, that combined intensive psychosocial therapy with vocational skills training over 12 months, using a randomized controlled trial with follow-up surveys at 1 year and at 3.5 years after program

completion. Three main findings emerge. First, the program produced large and statistically significant reductions in emotional, physical, and physical/sexual IPV at Endline 1, with magnitudes of 27–42% of control prevalence; these effects survive multiple-testing adjustment at conventional thresholds. The treatment group's IPV levels are approximately flat between Endline 1 and Endline 2; the control group converges downward. Second, the economic arm of the program, whose starter-kit grants were not delivered, produced only a short-run expenditure effect and no detectable effects on labor supply, earnings, or business ownership at either endline. Third, measured psychological well-being effects are null at both endlines, but qualitative evidence from focus group discussions conducted 3.5 years after the program suggests that treatment women continued to apply psychosocial skills (anger management, de-escalation, communication, self-efficacy) learned during the curriculum.

Our preferred reading of the Endline 2 convergence is not that the program's effects faded; the treatment group's IPV levels remained approximately flat, while the control group converged downward. We interpret this within the ecological framework developed in Section 4: the program appears to have produced durable change at the individual and interpersonal levels, where its curriculum operated most directly, while the concurrent scale-up of community- and societal-level interventions in Greater Monrovia plausibly affected both treatment and control. The one longer-term outcome that does not converge, perceived community norms against IPV (still 0.24 SD lower in the treatment group at 3.5 years), is consistent with this interpretation: the program shifted an interpretive frame that persisted independently of short-run individual attitudes and independently of whether the capital component was delivered.

The paper provides one of the longer post-program evaluations to date of a multifaceted female empowerment program.¹¹ This paper also documents IPV reductions sustained over 3.5 years in the treatment group at magnitudes broadly comparable to the non-economics meta-analytic literature on psychosocial IPV interventions, even though the capital-grant component of the program was not delivered. Moreover, through a second endline combined with contextual policy data, it demonstrates that short-run treatment effects can appear to “fade” at a longer-term follow-up not because they decay in the treated group, but because the counterfactual moves. This distinction has direct implications for how the policy literature interprets longer-term IPV follow-up evaluations.

Several limits are worth repeating. The sample is modest (less than 400 individual clusters), which bounds what can be learned from non-significant estimates. The single implemented cohort, the non-delivery of starter-kit grants, and the COVID-era data-collection conditions all shape external validity. We cannot causally attribute the control-side reduction in IPV to a specific contemporaneous program; the evidence is circumstantial, drawing on the concentration of policy activity in the study area and on the absence of aggregate labor-market or price-level changes that would otherwise explain the pattern. And the FGDs provide qualitative, not necessarily causal, evidence. Future work integrating administrative data on contemporaneous policy programs, and longer post-program survey waves across multiple comparable contexts, would help clarify when, and through which ecological levels, bundled empowerment programs produce durable effects on IPV.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

¹¹ For comparison, follow-up windows in the most-cited multifaceted-empowerment evaluations are: Bandiera et al. (2020), BRAC ELA in Uganda, four years post-program among adolescent girls; Abramsky et al. (2014), SASA! in Uganda, approximately three years; Green et al. (2015), two to four years post-program in northern Uganda; most psychosocial IPV trials reviewed in Turner et al. (2020) follow at twelve months or less.

Table A.1

WIN program components.

Program component	Description
Psychological support	One-to-one and group counseling, stress management, family/couple therapy
Literacy classes	Reading and writing curriculum by Ministry of Education
Child care	During program participation
Medical checkups	Free primary medical check-ups at Red Cross clinic
Vocational skills training	Baking, cosmetology, and tailoring
Entrepreneurship training	Financial literacy, business planning/management, etc.
Business start-up capital	250 USD worth of capital along with 30 USD cash grant

Table A.2

Selection criteria of WIN program.

1. Ex-combatant	5. Single mother/self-supported
2. Previous commercial sex worker	6. Illiterate
3. Victims of rape/domestic violence	7. Economically vulnerable
4. Witness of extreme violence	8. Drug user

Appendix A

See [Tables A.1–A.9](#) and [Figs. A.1](#) and [A.2](#).

Appendix B. Survey instrument

Controlling behavior

1. Did your man ever try to keep you from seeing your friends in the past 12 months?
2. Did your man ever try to stop you from meeting or speaking to your family of birth in the past 12 months?

3. Did your man ever need to know where you are all the time in the past 12 months?
4. Did your man ever stop talking to you or treat you with no interest in the past 12 months?
5. Did your man ever get angry if you speak with another man in the past 12 months?
6. Did your man often think that you are unfaithful in the past 12 months?
7. In the past 12 months, did your man ever expect you to ask for his approval before you go to a health clinic or hospital?

Emotional IPV¹²

1. Did your man ever insult you or make you feel bad about yourself in the past 12 months?

¹² For each ipv question, if the answer is “yes”, a follow-up question about frequency appears, asking whether it happened (i) one or two times, (ii) three to five times, or (iii) more than five times.

Table A.3

Attrition balance.

	(1) Endline 1 (April 2021) = 1 if completed endline survey	(2) = 1 if completed IPV module	(3) Endline 2 (Dec 2023–Jan 2024) = 1 if completed endline survey	(4) = 1 if completed IPV module
Treatment	0.00 (0.03)	−0.02 (0.04)	0.03 (0.05)	0.01 (0.05)
Control mean	0.91	0.81	0.70	0.62
Overall mean	0.91	0.79	0.72	0.62
Observations	395	395	395	395

Note: Sample is the N = 395 baseline cohort. Columns (1)–(2) report Endline 1 (April 2021); columns (3)–(4) report Endline 2 (December 2023–January 2024). Columns (1) and (3) report a binary indicator for whether the respondent was reached and completed the follow-up survey; columns (2) and (4) report a binary indicator for whether the respondent was eligible for the IPV module.^a Each column reports the treatment–control difference from a regression on the treatment indicator. Regressions include strata fixed effects. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

^a The IPV questionnaire is administered only to women who are currently partnered or have been partnered in the 12 months prior to the survey.

Table A.4

Frequency-weighted IPV measures.

	(1) IPV frequency sum	(2) Any IPV (binary)
Panel A. 1st Follow-up (1 year post program end)		
WIN treatment	−1.47 (1.09)	−0.14** (0.07)
Control mean	5.91	0.66
Control SD	7.79	0.48
Observations	169	169
Panel B. 2nd Follow-up (3.5 years post program end)		
WIN treatment	−0.12 (0.94)	0.01 (0.06)
Control mean	3.98	0.47
Control SD	7.56	0.50
Observations	245	245

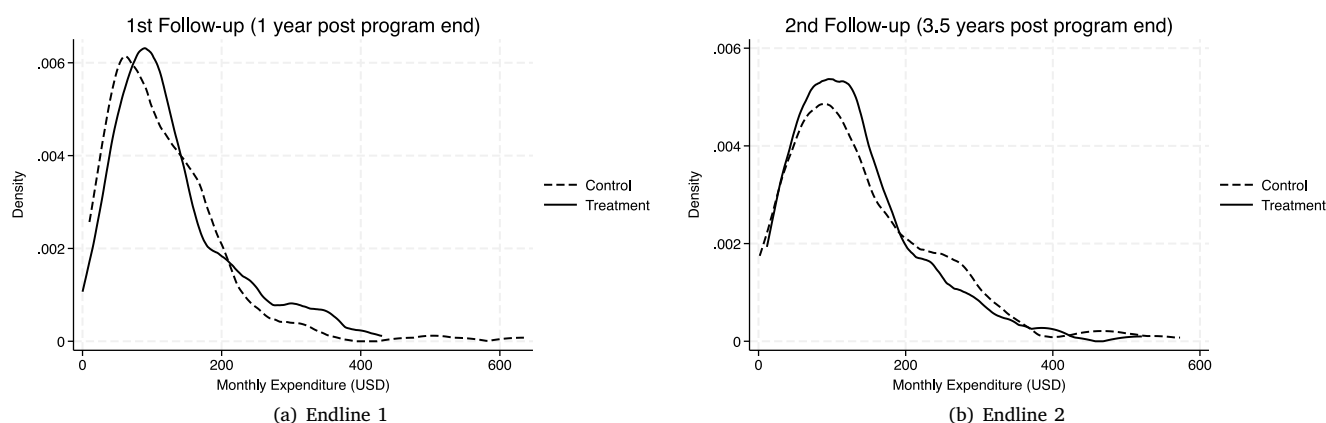
Note: Each outcome is the sum of the yes/no IPV indicators weighted by the reported frequency code. Specification and controls match [Table 2](#). Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.5

Robustness: Main IPV effects controlling for age and number of children.

	(1) Controlling behavior	(2) Emotional violence	(3) Physical violence	(4) Sexual violence	(5) Any IPV
Panel A. 1st Follow-up — with age and children controls					
WIN treatment	−0.02 (0.07)	−0.17** (0.07)	−0.18** (0.07)	−0.09 (0.07)	−0.15** (0.07)
Sharpened <i>q</i> -value	0.717	0.055	0.055	0.205	0.057
Control mean	0.80	0.62	0.45	0.24	0.66
Observations	169	169	169	169	169
Panel B. 2nd Follow-up — with age and children controls					
WIN treatment	−0.05 (0.05)	−0.01 (0.06)	0.00 (0.06)	−0.00 (0.04)	0.01 (0.06)
Sharpened <i>q</i> -value	0.999	0.999	0.999	0.999	0.999
Control mean	0.80	0.46	0.23	0.12	0.47
Observations	245	245	245	245	245

Note: Specification augments Table 2 with controls for respondent age and number of children at baseline. Standard errors are heteroskedasticity-robust. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

**Fig. A.1.** Expenditure distributions by treatment status.

Note: Kernel density estimates of household expenditures (USD) for treatment and control women. Panel (a) is Endline 1 (April 2021, N=359); Panel (b) is Endline 2 (December 2023–January 2024, N=283). Recall period and unit follow those in Table 7.

2. Did your man ever make you feel small in front of other people in the past 12 months?
3. Did your man ever mean to scare you (for example, by the way he looked at you, by yelling and bursting things) in the past 12 months?
4. Did your man ever threaten to hurt you or someone you care about in the past 12 months?

Physical IPV¹²

1. Did your man ever slap you or throw something at you that could hurt you in the past 12 months?
2. Did your man ever push you, shove you, or pull your hair in the past 12 months?
3. Did your man ever hit you with his hand or with something else that could hurt you in the past 12 months?
4. Did your man ever kick you, drag you or beat you up in the past 12 months?
5. Did your man ever mean to choke or burn you in the past 12 months?
6. Did your man ever threaten to use or actually use a gun, knife or other weapon against you in the past 12 months?

Sexual IPV¹²

1. Did your man ever physically force you to do man and woman business when you did not want to in the past 12 months?

2. Did you ever do man and woman business when you did not want to because you were afraid of what your man might do in the past 12 months?
3. In the past 12 months, while doing man and woman business, did your man ever force you to do something that made you feel small or bad about yourself?

Non-sensitive placebo questions

1. Did it rain in your village one time or more in the past year?
2. Did you do any farm work in the past year?
3. Did you sleep in the past week, during day or night?
4. Did you go to the market in the past week?
5. Did you travel outside of Liberia in the past week?
6. Will you, or anyone in your household, eat any rice next week, one time or more?
7. Will you, or anyone in your household, eat any type of meat next week, one time or more?

Data availability

Data will be made available on request.

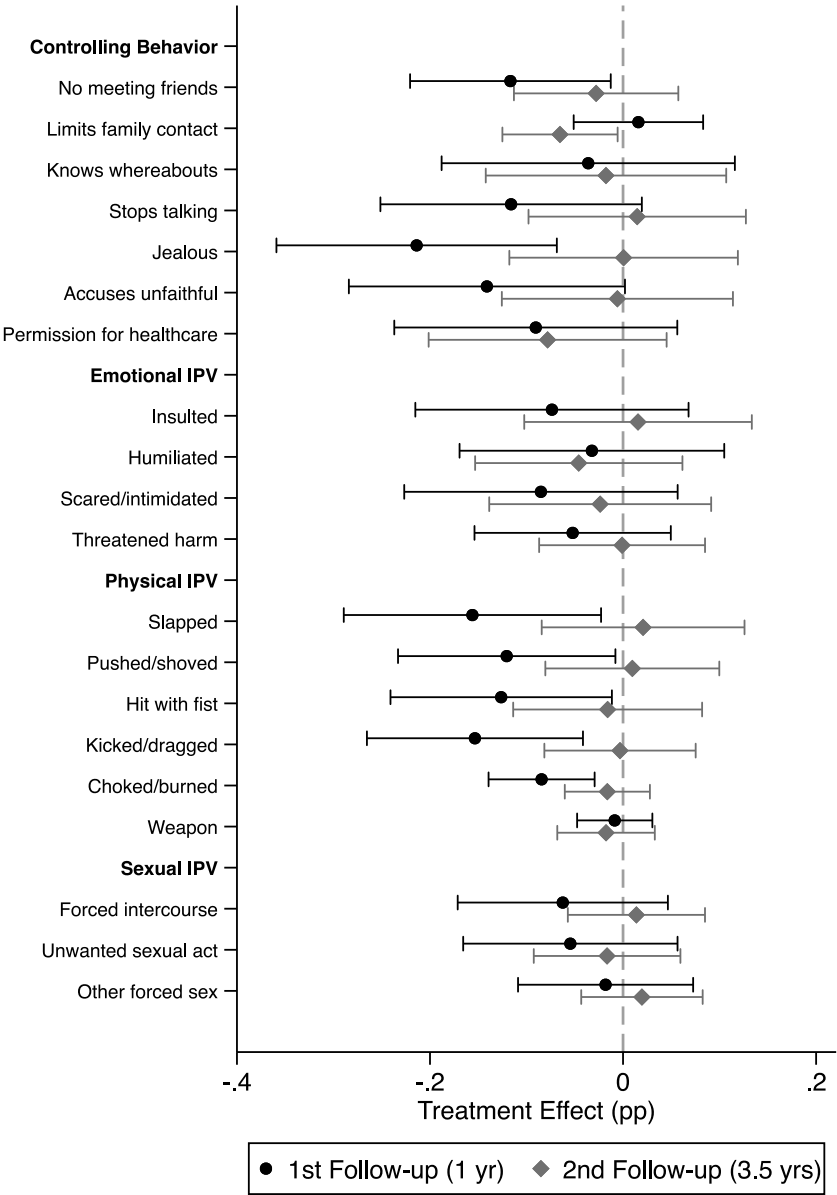


Fig. A.2. Treatment effects on individual IPV questions (coefficient plot).
Note: Each marker is the ANCOVA treatment-effect estimate for one individual IPV question (see Table A.6), with 95% confidence intervals based on heteroskedasticity-robust standard errors. Specification matches Table 2. Endline 1 (April 2021, FTFI arm only, N=169) and Endline 2 (December 2023–January 2024, all FTFI, N=245).

Table A.6
Treatment effects on individual IPV questions.

	1st Follow-up (1 year)		2nd Follow-up (3.5 years)	
	Treatment effect	Control mean	Treatment effect	Control mean
<i>Controlling Behavior</i>				
No meeting friends	−0.117** (0.053) [q = 0.091]	0.205	−0.028 (0.043) [q = 0.993]	0.148
Limits family contact	0.016 (0.034) [q = 0.592]	0.045	−0.065** (0.030) [q = 0.660]	0.090
Knows whereabouts	−0.036 (0.077) [q = 0.592]	0.598	−0.018 (0.064) [q = 0.993]	0.582
Stops talking	−0.116* (0.069) [q = 0.180]	0.345	0.015 (0.057) [q = 0.993]	0.262
Jealous	−0.214*** (0.074) [q = 0.038]	0.568	0.001 (0.060) [q = 0.993]	0.328
Accuses unfaithful	−0.141* (0.073) [q = 0.117]	0.500	−0.006 (0.061) [q = 0.993]	0.320
Permission for healthcare	−0.090 (0.075) [q = 0.372]	0.398	−0.078 (0.063) [q = 0.993]	0.434
<i>Emotional IPV</i>				
Insulted	−0.074 (0.072) [q = 0.381]	0.500	0.016 (0.060) [q = 0.993]	0.328
Humiliated	−0.032 (0.070) [q = 0.592]	0.364	−0.046 (0.055) [q = 0.993]	0.254
Scared/intimidated	−0.085 (0.072) [q = 0.372]	0.432	−0.024 (0.059) [q = 0.993]	0.328
Threatened harm	−0.052 (0.052) [q = 0.381]	0.170	−0.001 (0.044) [q = 0.993]	0.131
<i>Physical IPV</i>				
Slapped	−0.156** (0.068) [q = 0.091]	0.386	0.021 (0.054) [q = 0.993]	0.205
Pushed/shoved	−0.121** (0.057) [q = 0.091]	0.250	0.010 (0.046) [q = 0.993]	0.140
Hit with fist	−0.126** (0.059) [q = 0.091]	0.261	−0.016 (0.050) [q = 0.993]	0.189
Kicked/dragged	−0.153*** (0.057) [q = 0.045]	0.273	−0.003 (0.040) [q = 0.993]	0.107
Choked/burned	−0.084*** (0.028) [q = 0.038]	0.092	−0.016 (0.022) [q = 0.993]	0.041
Weapon	−0.009 (0.020) [q = 0.592]	0.023	−0.018 (0.026) [q = 0.993]	0.049
<i>Sexual IPV</i>				
Forced intercourse	−0.062 (0.056) [q = 0.372]	0.170	0.014 (0.036) [q = 0.993]	0.075
Unwanted sexual act	−0.055 (0.057) [q = 0.381]	0.182	−0.017 (0.039) [q = 0.993]	0.098
Other forced sex	−0.018 (0.046) [q = 0.592]	0.102	0.020 (0.032) [q = 0.993]	0.057

Note: Each row is a separate ANCOVA regression for one of the individual IPV questions in the WHO Violence Against Women module. Specification matches Table 2. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.7

Heterogeneous treatment effects on IPV by baseline physical/sexual IPV status.

	(1) Controlling behavior	(2) Emotional violence	(3) Physical violence	(4) Sexual violence	(5) Any IPV
Panel A. 1st Follow-up — HTE by Baseline Physical/Sexual IPV					
WIN treatment	−0.03 (0.09)	−0.30*** (0.10)	−0.27*** (0.09)	−0.21*** (0.07)	−0.28*** (0.10)
Treatment × Baseline Phys/Sex IPV	0.02 (0.13)	0.28* (0.14)	0.16 (0.14)	0.28** (0.13)	0.27* (0.14)
Control mean	0.80	0.62	0.45	0.24	0.66
Observations	169	169	169	169	169
Panel B. 2nd Follow-up — HTE by Baseline Physical/Sexual IPV					
WIN treatment	−0.04 (0.07)	−0.02 (0.08)	0.01 (0.07)	−0.02 (0.05)	0.01 (0.08)
Treatment × Baseline Phys/Sex IPV	−0.04 (0.11)	0.02 (0.13)	0.02 (0.11)	0.07 (0.09)	0.00 (0.12)
Control mean	0.80	0.46	0.23	0.12	0.47
Observations	245	245	245	245	245

Note: Specification augments Table 2 with a treatment × baseline-physical/sexual-IPV interaction. Baseline physical/sexual IPV status is the pre-registered stratification variable. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.8

Heterogeneous treatment effects on IPV by baseline civil-war exposure.

	(1) Controlling behavior	(2) Emotional violence	(3) Physical violence	(4) Sexual violence	(5) Any IPV
Panel A. 1st Follow-up — HTE by Civil War Exposure					
WIN treatment	−0.08 (0.09)	−0.14 (0.10)	−0.08 (0.10)	−0.10 (0.08)	−0.10 (0.10)
Treatment × Civil War Affected	0.13 (0.13)	−0.05 (0.14)	−0.25* (0.15)	0.06 (0.13)	−0.11 (0.14)
Control mean	0.80	0.62	0.45	0.24	0.66
Observations	169	169	169	169	169
Panel B. 2nd Follow-up — HTE by Civil War Exposure					
WIN treatment	−0.10 (0.07)	−0.07 (0.08)	0.00 (0.07)	−0.06 (0.06)	−0.04 (0.08)
Treatment × Civil War Affected	0.09 (0.11)	0.15 (0.12)	0.02 (0.11)	0.15* (0.08)	0.11 (0.12)
Control mean	0.80	0.46	0.23	0.12	0.47
Observations	245	245	245	245	245

Note: Specification augments Table 2 with a treatment × baseline-civil-war-exposure interaction. Baseline civil-war exposure is the second pre-registered stratification variable. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.9

Robustness: Labor supply estimates with hours winsorized at the 5th and 95th percentiles.

	(1) Self-empl. hours	(2) Casual hours	(3) Other hours	(4) Total hours
Panel A. 1st Follow-up — Winsorized at 5/95				
WIN treatment	13.14 (8.39)	−0.09 (0.11)	−0.36 (0.29)	10.63 (8.59)
Control mean	35.87	0.31	1.06	46.64
Observations	359	359	359	359
Panel B. 2nd Follow-up — Winsorized at 5/95				
WIN treatment	−8.09 (8.97)	0.14 (0.17)	−0.61* (0.33)	−15.28* (9.03)
Control mean	42.98	0.41	1.27	53.97
Observations	283	283	283	283

Note: Specification matches the labor-supply regressions in the main text after symmetrically winsorizing monthly hours at the 5th and 95th percentiles. Heteroskedasticity-robust standard errors in parentheses. Significance levels: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

References

- Abramsky, Tanya, Devries, Karen, Kiss, Ligia, Nakuti, Janet, Kyegombe, Nambusi, Starmann, Elizabeth, Cundill, Bonnie, Francisco, Leilani, Kaye, Dan, Musuya, Tina, Michau, Lori, Watts, Charlotte, 2014. Findings from the SASA! study: A cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. *BMC Med.* 12, 122.
- Alfonsi, Livia, Bandiera, Oriana, Bassi, Vittorio, Burgess, Robin, Rasul, Imran, Sulaiman, Munshi, Vitali, Anna, 2020. Tackling youth unemployment: Evidence from a labor market experiment in Uganda. *Econometrica* 88 (6), 2369–2414.
- Anderson, Michael L., 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *J. Amer. Statist. Assoc.* 103 (484), 1481–1495.
- Angelucci, Manuela, 2008. Love on the rocks: Domestic violence and alcohol abuse in rural Mexico. *B.E. J. Econ. Anal. Policy* 8 (1).
- Attanasio, Orazio, Guarín, Arlen, Medina, Carlos, Meghir, Costas, 2017. Vocational training for disadvantaged youth in Colombia: A long-term follow-up. *Am. Econ. J.: Appl. Econ.* 9 (2), 131–143.
- Bachus, Loraine J., Ranganathan, Meghna, Watts, Charlotte, Devries, Karen, 2018. Recent intimate partner violence against women and health: a systematic review and meta-analysis of cohort studies. *BMJ Open* 8 (7), e019995.
- Bandiera, Oriana, Buehren, Niklas, Burgess, Robin, Goldstein, Markus, Gulesci, Selim, Rasul, Imran, Sulaiman, Munshi, 2020. Women's empowerment in action: Evidence from a randomized control trial in Africa. *Am. Econ. J.: Appl. Econ.* 12 (1), 210–259.
- Bandiera, Oriana, Burgess, Robin, Das, Narayan, Gulesci, Selim, Rasul, Imran, Sulaiman, Munshi, 2017. Labor markets and poverty in village economies. *Q. J. Econ.* 132 (2), 811–870.
- Benjamini, Yoav, Krieger, Abba M., Yekutieli, Daniel, 2006. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* 93 (3), 491–507.
- Blattman, Christopher, Fiala, Nathan, Martinez, Sebastian, 2014. Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *Q. J. Econ.* 129 (2), 697–752.
- Blattman, Christopher, Green, Eric P., Jamison, Julian, Lehmann, M. Christian, Annan, Jeannie, 2016. The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda. *Am. Econ. J.: Appl. Econ.* 8 (2), 35–64.
- Blattman, Christopher, Jamison, Julian C., Sheridan, Margaret, 2017. Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia. *Am. Econ. Rev.* 107 (4), 1165–1206.
- Blattman, Christopher, Ralston, Laura, 2015. Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs. <http://dx.doi.org/10.2139/ssrn.2622220>.
- Bobonis, Gustavo J., Gonzalez-Brenes, Melissa, Castro, Roberto, 2013. Public transfers and domestic violence: The roles of private information and spousal control. *Am. Econ. J.: Econ. Policy* 5 (1), 179–205.
- Buller, Ana Maria, Peterman, Amber, Ranganathan, Meghna, Bleile, Alexandra, Hidrobo, Melissa, Heise, Lori, 2018. A mixed-method review of cash transfers and intimate partner violence in low- and middle-income countries. *World Bank Res. Obs.* 33 (2), 218–258.
- Domingo, Pilar, Holmes, Rebecca, Menocal, Alina Rocha, Jones, Nicola, 2013. Assessment of the Evidence of Links Between Gender Equality, Peacebuilding and Statebuilding. ODI (Overseas Development Institute) Report.
- Ellsberg, Mary, Arango, Diana J., Morton, Matthew, Gennari, Floriza, Kiplesund, Sveinung, Contreras, Manuel, Watts, Charlotte, 2015. Prevention of violence against women and girls: what does the evidence say? *Lancet* 385 (9977), 1555–1566.
- Green, Eric P., Blattman, Christopher, Jamison, Julian, Annan, Jeannie, 2015. Women's entrepreneurship and intimate partner violence: A cluster randomized trial of microenterprise assistance and partner participation in post-conflict Uganda. *Soc. Sci. Med.* 133, 177–188.
- Haushofer, Johannes, Mudida, Robert, Shapiro, Jeremy P., 2020. The Comparative Impact of Cash Transfers and a Psychotherapy Program on Psychological and Economic Well-being. (w28106), National Bureau of Economic Research.
- Haushofer, Johannes, Ringdal, Charlotte, Shapiro, Jeremy P., Wang, Xiao Yu, 2019. Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya. National Bureau of Economic Research, Working Paper 25627.
- Heath, Rachel, Hidrobo, Melissa, Roy, Shalini, 2020. Cash transfers, polygamy, and intimate partner violence: Experimental evidence from Mali. *J. Dev. Econ.* 143, 102410.
- Heise, Lori, 1998. Violence against women: An integrated, ecological framework. *Violence Against Women* 4 (3), 262–290.
- Hidrobo, Melissa, Fernald, Lia, 2013. Cash transfers and domestic violence. *J. Health Econ.* 32 (1), 304–319.
- Hidrobo, Melissa, Peterman, Amber, Heise, Lori, 2016. The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern Ecuador. *Am. Econ. J.: Appl. Econ.* 8 (3), 284–303.
- Jones, Nicola, Cooper, Janice, Presler-Marshall, Elizabeth, Walker, David, 2014. The Fallout of Rape as a Weapon of War. ODI (Overseas Development Institute) Report.
- Kelly, Jocelyn T.D., Colantuoni, Elizabeth, Robinson, Courtland, Decker, Michele R., 2021. Quantifying the ripple effects of civil war: How armed conflict is associated with more severe violence in the home. *Health Hum. Rights* 23 (1), 75–89.
- Kelly, Jocelyn, Ekhat-Mobayode, Uche, Hanmer, Lucia C., Rubin, Amalia, Arango, Diana J., 2024. The risk that travels with you: Links between forced displacement, conflict and intimate partner violence in Colombia and Liberia. *J. Dev. Stud.* 60 (12), 1994–2021.
- McKenzie, David, Puerto, Susana, 2021. Growing markets through business training for female entrepreneurs: a market-level randomized experiment in Kenya. *Am. Econ. J.: Appl. Econ.* 13 (2), 297–332.
- Moser, Caroline, Winton, Ailsa, 2002. Violence in the central American region: Towards an integrated framework for violence reduction. ODI (Overseas Development Institute) Working Paper.
- Omanyondo, Marie-Claire, 2005. Sexual gender-based violence and health facility needs assessment. WHO Report.
- Park, David Sungho, 2020. Reducing intimate partner violence of disadvantaged women in Monrovia: An experiment with red cross's women training and integration program. <http://dx.doi.org/10.1257/rct.4488>, AEA RCT Registry, AEARCTR-0004488.
- Park, David Sungho, Aggarwal, Shilpa, Jeong, Dahyeon, Kumar, Naresh, Robinson, Jonathan, Spearot, Alan, 2025. Private but misunderstood? Evidence on measuring intimate partner violence via self-interviewing in rural Liberia and Malawi. *World Bank Econ. Rev.* 39 (4), 907–923.
- Peterson, Cora, Kearns, Megan C., McIntosh, Wendy LiKamWa, Estefan, Lianne Fuino, Nicolaidis, Christina, McCollister, Kathryn E., Gordon, Amy, Florence, Curtis, 2018. Lifetime economic burden of intimate partner violence among U.S. adults. *Am. J. Prev. Med.* 55 (4), 433–444.
- Ranganathan, Meghna, Heise, Lori, Peterman, Amber, Roy, Shalini, Hidrobo, Melissa, 2021. Cross-disciplinary intersections between public health and economics in intimate partner violence research. *SSM - Popul. Health* 14, 100822.
- Roy, Shalini, Hidrobo, Melissa, Hoddinott, John, Ahmed, Akhter, 2019. Transfers, behavior change communication, and intimate partner violence: Postprogram evidence from rural Bangladesh. *Rev. Econ. Stat.* 101 (5), 865–877.
- Sardinha, Lynnmarie, Maheu-Giroux, Mathieu, Stöckl, Heidi, Meyer, Sarah Rachel, García-Moreno, Claudia, 2022. Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *Lancet* 399 (10327), 803–813.
- Smith, S.G., Chen, J., Basile, K.C., Gilbert, L.K., Merrick, M.T., Patel, N., Walling, M., Jain, A., 2017. The national intimate partner and sexual violence survey (NISVS): 2010–2012 state report.
- Steenkamp, Chrissie, 2005. The legacy of war: Conceptualizing a 'culture of violence' to explain violence after peace accords. *Round Table* 94 (379), 253–267.
- Turner, Daniel T., Riedel, Eva, Kobeissi, Loulou H., Karyotaki, Eirini, Garcia-Moreno, Claudia, Say, Lale, Cuijpers, Pim, 2020. Psychosocial interventions for intimate partner violence in low and middle income countries: A meta-analysis of randomised controlled trials. *J. Glob. Health* 10 (1), 010409.
- UN Women, 2013. The contribution of UN women to increasing women's leadership and participation in peace and security and in humanitarian response. UN Women Evaluation Report.
- WHO, 2016. Ethical and safety recommendations for intervention research on violence against women.
- WHO, 2024. Violence against Women. World Health Organization, URL <https://www.who.int/news-room/fact-sheets/detail/violence-against-women>.