Policy Research Working Paper 11040

The Unintended Impacts of an Intimate Partner Violence Prevention Program

Experimental Evidence from Rwanda

Claire Cullen
Arthur Alik-Lagrange
Mũthoni Ngatia
Julia Vaillant

WORLD BANK GROUP
Africa Region
Gender Innovation Lab
January 2025

Policy Research Working Paper 11040

Abstract

This study evaluates the impact of an intimate partner violence prevention program in Rwanda, using a randomized controlled trial. The 22-week couples training program aimed to improve communication, shift gender attitudes, and promote gender equality. Randomizing at both the village and couple levels, the study finds that the program caused large, unexpected increases in intimate partner violence. Treated women reported 5 and 10 percentage points more physical and sexual violence than control women, respectively, and control women living in treated villages reported even larger increases of 11 and 17 percentage points. Evidence suggests that these increases stem from male backlash against perceived identity threats, as women's more progressive attitudes and aspirations contrast with men's pressure to maintain traditional norms, generating friction within households. Innovative measurement techniques and extensive robustness checks show that changes in reporting behavior cannot explain the results.

This paper is a product of the Gender Innovation Lab, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at jvaillant@worldbank.org.

The Unintended Impacts of an Intimate Partner Violence Prevention Program: Experimental Evidence from Rwanda

Claire Cullen, Arthur Alik-Lagrange, Mũthoni Ngatia, Julia Vaillant*

***JEL*: C93, J12, J16, O12**

Keywords: Intimate Partner Violence, Gender, Identity, Spillovers, Norms, Field Experiments

*Cullen (claire.cullen010@gmail.com), Alik-Lagrange (aaliklagrange@worldbank.org), Ngatia (ngatia@gmail.com), Vaillant (jvaillant@worldbank.org). Acknowledgments: We thank Julien Labonne, Markus Goldstein, and Rachael Pierotti for support throughout the project. We thank the respondents for participating in the surveys, and the staff at Laterite Ltd for their outstanding data collection work. We also thank Jean Paul Ngabonziza, Victoria Isika, Harald Hinkel, Verena Phipps, Patricia Fernandes, Jessica Ott, Stephanie Stahlberg, Ludovic Bequet, and Alex Twahirwa and other staff at MIGEPROF, and CARE, RWAMREC, and Rwandan Women's Network for all their support, programming, research assistance, or insights into program administration. We are grateful to Lori Heise, Kristin Dunkle, Erin Stern, and other members of the Indashyikirwa team for their roles in the original program and for fruitful discussions of the results. We are grateful for comments and suggestions from Noam Angrist, Diana Arango, Leonardo Burzstyn, Cheryl Doss, Eliana La Ferrara, Anandi Mani, Ted Miguel, Berk Özler, Simon Quinn, Maria Micaela Sviatschi, and seminar participants at NEUDC, Oxford University, Princeton ECBE, CSAE, Melbourne University, Australian Gender Economics Workshop, the World Bank, and SVRI. All errors are our own. The study protocols were approved by the Rwandan National Ethics Committee (722/RNEC/2018) and the Oxford University Central University Research Ethics Committee (R50347), and a research visa was granted by the National Institute of Statistics Rwanda (0676/2018/10/NISR). This RCT was registered as AEARCTR-0002282. This paper is a product of the World Bank Africa Gender Innovation Lab, within the Office of the Africa Region Chief Economist. We are grateful to the World Bank's Umbrella Facility for Gender Equality (UFGE) for financial support. The UFGE is a multi-donor trust fund administered by the World Bank to advance gender equality and women's empowerment, and is supported by Australia, Canada, Denmark, Finland, Germany, Iceland, Ireland, Latvia, the Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom, United States, the Bill and Melinda Gates Foundation, and the Wellspring Philanthropic Fund. We are also grateful for support from the World Bank's Nordic Trust Fund and the Swiss Development Cooperation. The findings and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the Government of Rwanda, of the World Bank and its affiliated organizations, or those of its Executive Directors or the governments they represent.

1 Introduction

Intimate partner violence (IPV) affects one in three women in their lifetime (García-Moreno et al., 2015), with some of the highest IPV rates found in Sub-Saharan Africa (Cools and Kotsadam, 2017). Aside from being a human rights issue, IPV is also estimated to cost between 1.2 and 5 percent of global gross domestic product (GDP) (Fearon and Hoeffler, 2014; Klugman et al., 2014). These costs include negative impacts on women's labor force participation, and women's and children's physical and mental health. These effects can be passed on through generations, with the children of IPV victims having lower birth weight and IQ, worse school performance, and being more likely to enter abusive relationships as adults (Aizer, 2011; Carrell et al., 2018; Roy et al., 2019).

The persistence of IPV and other gender gaps across all economies and at all income levels suggests that gender equality does not automatically improve as countries become richer. Instead, sticky gender identities and social norms appear to play a role in maintaining gender inequality and may be vital to tackle to overcome high rates of IPV. In spite of the scale of the issue, the evidence on what works to reduce IPV is still limited. Further, policies designed to change underlying gender norms or to promote gender equality may risk provoking a backlash, exacerbating the issue they sought to address (Anderson and Genicot, 2015; Bhalotra et al., 2021; Cullen et al., 2024; Erten and Keskin, 2018).

In this paper, we present the results of one of the first large-scale randomized controlled trials (RCTs) of a government-managed couples-based IPV prevention program. The intervention was a group-based training program for couples led by the Rwandan Ministry of Gender and Family Promotion and implemented by local and international nongovernmental organizations (NGOs). The program, developed for the Rwandan context, consisted of 22 weekly facilitated group discussions with an average of 19 couples per group. It was designed to prevent IPV by improving couples' marital communication skills, changing conservative gender attitudes, and promoting new progressive gender norms in the community. Trained facilitators delivered information and skills through structured sessions including group activities and discussions, and couples' take-home exercises.

Using a two-level randomized trial design, 98 villages and 2,042 couples belonging to village savings and loans associations (VSLAs) were assigned to treatment or control groups, producing

three study groups: a treatment group, a spillover group, and a control group.[1] This design allows us to identify both the direct and indirect effects of the intervention by comparing the treated and the spillover groups with the control group, respectively. Six months post-program, we measured impacts on women's self-reported experience of emotional, physical, and sexual IPV, and a range of other indicators including gender attitudes, well-being, and economic empowerment.

We find large, unintended, negative impacts of the IPV prevention program on both treated and spillover couples. Relative to pure control couples, women who attended the couples training report more emotional, physical, and sexual IPV (7, 5, and 10 percentage points more, respectively). Spillover couples have even larger increases, of 7, 11, and 17 percentage points, respectively. Negative IPV impacts are also reflected in broadly negative results across other wellbeing indicators, particularly for spillover couples. For example, depression, alcohol consumption, and IPV against men also increase.

Evidence on mechanisms suggests that by empowering women in the community, the program may have threatened men's sense of identity and provoked a widespread IPV backlash. Results are consistent with this backlash being driven by increased intra-household friction due to a divergence in attitudes and desire for power and agency within couples. Women become more progressive and desire more input in decision-making. Meanwhile, husbands either become more conservative, or husbands' slightly improved gender attitudes are overtaken by the much larger progressive shift in their wives' attitudes. Results suggest this generates tension within couples (e.g., increased cheating and escalation of disagreements to violence, and worsened relationship quality). Additionally, qualitative reports indicate that the diffusion of progressive messages across the community leads to shaming and pressure on men to reassert traditional male identities through violence.

The primary outcomes of women's experience of emotional, physical, and sexual IPV are self-reported. Given the sensitive nature of the topic of IPV and the possible correlation between a progressive gender norm–changing program and IPV reporting, there is a concern that treatment effects might reflect changes in reporting rather than true increases in IPV. However, we present evidence from numerous robustness checks suggesting this is not the case. We conduct extensive analysis on many plausible reporting channels, and find that the data suggest effects are not driven

[1]Note that the 'spillover' group comprised control couples in treated villages that belong to the same VSLAs as treated couples.

by changes in reporting from: an expanded understanding of IPV questions; reduced stigma or empowerment to report; or strategically misreporting violence to get access to programs.[2] Our results are robust to the use of alternative specifications.

This study brings together two strands of the growing economic literature on the causes of IPV: the literature that has focused on women's relative bargaining power in the household, and literature on identity and cultural factors, such as norms, that affect economic behavior (Akerlof and Kranton, 2000; Field et al., 2021; Guarnieri and Tur-Prats, 2023; LaMattina, 2017; Oh, 2023). Most theoretical models of the economic behavior underlying IPV frame it as the result of non-cooperative intra-household bargaining between partners, where a husband's utility derived from violence is constrained by the wife's outside option (Bloch and Rao, 2002; Eswaran and Malhotra, 2011; Farmer and Tiefenthaler, 1997; Tauchen et al., 1991). IPV has also been presented as either instrumental — to extract pecuniary resources from the female partner — or expressive — where the husband receives non-pecuniary returns from violence; for example, if he enjoys asserting dominance (Haushofer et al., 2019). To date, most of the economic literature on IPV has focused on shifting bargaining power dynamics either from tangible material resources (such as cash transfers to wives) or changes in outside options that could affect material resources or stress levels (such as men's unemployment, eased divorce laws, or women's relative wages). The empirical literature primarily finds that boosting women's bargaining power tends to reduce or have no effect on IPV, on average (Baranov et al., 2020; Buller et al., 2018; J-PAL, 2022). Further, most economic models of IPV do not explain the underlying drivers of expressive violence, nor why expressive violence occurs if men do not have an implicit taste or preference for it.

This paper makes three key contributions to the literature. First, we contribute to the economic literature on IPV, providing some of the first evidence that violence can increase substantially even without changes in relative economic welfare between spouses. This finding advances our

[2]First, we rule out increased strategic reporting to maximize the likelihood of being 'selected' for future access to the program and its small travel stipend. This evidence includes, for example, all groups reporting an increase in men's alcohol consumption, and no impact on self-reported financial need. This is inconsistent with people trying to look like 'good' or 'in need' candidates. We also rule out reporting bias from an expanded understanding of what the IPV survey questions mean. For example, the IPV increase is smaller for treated women who spent 22 weeks discussing gender than for spillover women; and we find similar impacts across IPV questions that use more subjective and objective terms. Finally, we rule out women reporting more because of a reduction in the stigma of reporting. For example, we find almost no treatment effect on women's refusal-to-answer rate; and spillover women who experience the largest increase in IPV are no more likely to speak to others about the violence that they experienced, suggesting no reduced stigma. Finally, women's reports of experiencing violence are correlated with several non-self-reported indicators related to violence that have less risk of measurement bias, for example, questions answered by the enumerator about whether the woman had visible signs of injury or was referred to and spoke to a counselor after the survey. Further, we also find that spillover men report being more likely to perpetrate IPV in a set of hypothetical situations, corroborating their wives' reports of an IPV increase (though treated men report being less likely to perpetrate violence).

understanding beyond traditional bargaining models that emphasize material resources and outside options (Bloch and Rao, 2002; Farmer and Tiefenthaler, 1997; Tauchen et al., 1991) to highlight the critical role of identity threats. We build on the nascent literature examining non-material drivers of IPV (Adams et al., 2024; Tur-Prats, 2019) by connecting it to broader work showing how identity and social norms can shape behavioral and economic outcomes and responses to policy (Akerlof and Kranton, 2000; Bursztyn et al., 2020; Guarnieri and Tur-Prats, 2023; Karing, 2024; Oh, 2023). While previous studies have documented backlash against women's empowerment — whether through cash transfers, labor market opportunities, or legal reforms — these interventions primarily operated by shifting the material balance of power between spouses (Anderson and Genicot, 2015; Angelucci, 2008; Bhalotra et al., 2021; Bulte and Lensink, 2019; Erten and Keskin, 2021). Our results reveal that perceived threats to male identity alone can trigger significant increases in IPV - a finding that helps reconcile the mixed empirical evidence on IPV prevention programs with theoretical predictions that improvements in women's bargaining power can either increase or decrease violence (Eswaran and Malhotra, 2011; Tauchen et al., 1991).

Second, this study provides methodological insights for evaluating couples-based IPV prevention programs through a novel application of blinded, community-level randomization. While couples-based training programs are becoming a more common policy tool globally, existing evaluations face critical limitations that our design addresses. Several previous studies reporting positive or null effects used non-blinded recruitment that sometimes generated large baseline imbalances between treatment and control groups in key observable characteristics such as IPV (Dunkle et al., 2020; Sharma et al., 2020).[3] Other couples-training studies randomized beneficiaries within communities (Boyer et al., 2022; Doyle et al., 2018) but had no pure control group outside treated communities, potentially biasing treatment estimates through contamination. Our design overcomes these two sources of potential bias by implementing both geographic buffers between treatment and control villages and identical recruitment protocols across groups. The importance of this approach is evident in our results: we find larger IPV increases in spillover than treatment groups, meaning that without geographically separated controls, we would have (incorrectly) concluded the program *reduced* violence.[4] Other within-community individually randomized studies have found reductions

[3]For instance, Dunkle et al. (2020) found a difference of 10 percentage points in baseline IPV between treatment and control sectors.

[4]Consistent with our finding of negative spillovers, Chatterji et al. (2020) document negative spillover effects of *Indashyikirwa*

in IPV, but could not rule out that this was driven by increased IPV in the control group. This finding has broader implications for those evaluating programs targeting culturally sensitive norms, where community-level responses may dominate individual treatment effects.[5]

Beyond IPV, this study advances our understanding of how entrenched gender norms resist change. We contribute to a growing literature examining interventions aimed at shifting persistent gender gaps (Banerjee et al., 2019; Beaman et al., 2012; Dhar et al., 2022; Field et al., 2021), highlighting how attempts to alter deeply held attitudes and identities can trigger unintended consequences. Our findings suggest that while programs may successfully shift individual attitudes in the short term, community-level backlash can overwhelm these gains when interventions challenge traditional power structures. This speaks to fundamental questions in economics about how preferences and norms evolve, and when policy interventions may accelerate or impede social change (Becker, 2009; Benabou and Tirole, 2011; Giuliano and Nunn, 2021).

Couples-based IPV prevention programs similar to the one studied in this paper are currently being tested or scaled up in multiple countries around the world. Yet there is limited evidence of their effectiveness or potential spillover effects. Our results therefore have important policy implications for governments and other actors seeking to shift sensitive gender norms. They suggest caution should be exercised to mitigate the risks of backlash, and that alternative approaches to changing sensitive outcomes should be considered.

2 Study Context and Design

2.1 Country Context

Rwanda has experienced dramatic social change over the past 30 years, including a radical shift focused on women's status and empowerment. While modern national policy strongly promotes gender equality, this stands in tension with historically conservative gender norms that persist in many communities. Colonial rule and religious institutions reinforced patriarchal norms that limited women's autonomy and political power (Asiedu et al., 2018). These attitudes remain influential

on non-treated community members, with suggestive increases in sexual ($p<0.06$) and physical ($p<0.13$) IPV.

[5]The only other group-based IPV prevention study with both identical blind recruitment for treatment and control groups and geographic buffers was by Vaillant et al. (2020). Their 16-session program in the Democratic Republic of Congo worked with men's groups rather than couples and found no impact on IPV.

today. Rwanda's 2019–20 Demographic and Health Survey (DHS) reports that 40% of women report experiencing physical and/or sexual IPV in their lifetime, with 24% reporting incidents in the previous year.[6] Supporting norms remain widespread, particularly in rural areas, with 41% of women and 18% of men believing husband-to-wife violence can be justified in some circumstances (Mannell et al., 2018). This context is further shaped by the 1994 genocide's legacy of sexual violence, where rape was used as a weapon of war.

In parallel, over the past three decades, Rwanda has become a global leader in gender equality reforms. Following the genocide, when women comprised 70% of the population, women's organizations successfully advocated for female representation in the interim government (Hansén, 2017). Since then, Rwanda has enacted comprehensive reforms including equal land ownership rights (1999), a 30% gender quota for civil servants and parliamentarians (2003, now exceeded), and equal pay legislation (2009). The government also passed a Gender-Based Violence Bill requiring all citizens to report violence, and penalties of eight months to two years imprisonment for perpetrating IPV including 'harassing one's spouse' and marital rape (Mannell et al., 2018). At compulsory monthly community meetings, gender-based violence is regularly discussed. Recent reforms have overturned men's legal status as household heads (Abbott and Malunda, 2016), established IPV survivor support centers, and expanded women's rights in marriage, divorce, and inheritance — including protections for unofficially married women. These progressive legal reforms have been mirrored in health and education policies for women and girls, resulting in improved maternal mortality, contraceptive access, and girls' educational attainment (DHS, 2016).

However, a significant gap persists between national policy and local attitudes, particularly in rural areas. Traditional gender norms about women's subservience to men remain influential, and many are unaware of legal changes. For instance, in our study sample, only 24% of men and 12% of women know that Rwandan law establishes husband and wife as equal heads of household.

2.2 Couples Training Intervention

There is growing global interest in couples-based approaches to IPV prevention. These programs aim to reduce violence by engaging both partners in critical discussions about gendered power

[6]These rates increased from the 2014–15 DHS, which reported 34% lifetime IPV and 21% previous-year IPV. The DHS collects IPV data through face-to-face interviews.

dynamics and building skills in communication and conflict resolution. Through weeks of facilitated group activities and discussion and individual reflection, couples are encouraged to shift their attitudes and behaviors, and to model gender-equal, violence-free relationships in their communities. This approach is increasingly being tested and scaled by governments and NGOs across Sub-Saharan Africa, the Middle East, and Latin America. While earlier single-gender interventions have shown limited impact on IPV (Gupta et al., 2013; Hossain et al., 2014; Kerr-Wilson et al., 2020; Vaillant et al., 2020), couples-based programs may be more effective by simultaneously addressing men's motivation to change and fostering new relationship dynamics.

We evaluate a couples-based IPV prevention program led by Rwanda's Ministry of Gender Equality and Family Promotion (MIGEPROF). MIGEPROF worked with the NGOs CARE International, the Rwandan Women's Network, and Rwandan Men's Resource Centre (RWAMREC), and the World Bank provided funding to support the program. The program's central component was a five-month couples training, complemented by two elements available in both treatment and control communities (thus not evaluated in this study): a week-long training for local opinion leaders on gender equality, and safe spaces providing IPV support services.[7]

The couples training delivered 22 weekly group workshops over five months to 640 couples across 49 treated villages. Couples that were members of village savings and loan associations in 98 villages that had been selected for the study were invited to volunteer for the couples training. After the selection of 49 treatment villages, those households who were then randomly selected to participate met weekly for 3–4-hour discussions with an average of 18 other VSLA couples from their own or nearby treated villages. Each group was led by two trained facilitators, one male and one female, hired and trained by the Rwandan NGO, RWAMREC.

The sessions involved facilitated discussions and reflection on concepts including types and uses of power; gender and sex; rights; communication and conflict resolution skills; the benefits of non-violence and gender equality; managing triggers of violence including alcohol use; gender

[7]The intervention was based on two earlier programs, *Journeys of Transformation* (Slegh et al., 2013) and *SASA!* (Abramsky et al., 2014). These programs were adapted for the rural Rwandan context through extensive testing and piloting for the *Indashyikirwa* ('agents of change') program (Stern and Nyiratunga, 2017). *Indashyikirwa* comprised four components: couples training, community activism, opinion leader training, and safe spaces. To test a potentially more cost-effective model of the program, and to address the high demand from trained *Indashyikirwa* couples in attending the community activist training, several changes were made to the initial program: condensing the separate 10-day community activism curriculum into two additional sessions for all treated couples rather than to a select group of couples; adding a positive parenting session; and removing two cognitive skills restructuring sessions. The resulting program maintained 19 of 22 original *Indashyikirwa* sessions and was delivered by the same NGO consortium. The *Indashyikirwa* couples curriculum is available on https://www.whatworks.co.za/documents/publications/curricula/268-couple-curriculum-05-09-2018/file.

roles in the household; shared decision-making; and positive parenting. In later sessions, couples were encouraged to become 'agents of change' by supporting IPV survivors and modeling gender-equitable relationships in their communities. Each session included take-home activities for practice and reflection. The full curriculum is detailed in table A1. Participants received a nominal stipend of RWF 2,500 (approximately USD 2) per session to cover attendance costs, an amount likely exceeding participation costs given low formal employment in rural areas. An independent survey company monitored attendance during each session.

2.3 A Conceptual Framework of Identity and Intimate Partner Violence

We present a framework for understanding how gender identity can influence IPV and why programs targeting gender norms may inadvertently increase IPV. Traditional economic models of intimate partner violence emphasize bargaining power and outside options, focusing on how men use violence either instrumentally (to extract resources or coerce women to stay) or expressively due to inherent preferences (Bloch and Rao, 2002; Tauchen et al., 1991). However, the identity model suggests that individuals derive utility not only from material outcomes, but also from conforming to social categories and their associated prescriptions. When these identities are threatened — whether through one's own actions or others' behavior — individuals may engage in destructive behaviors to reaffirm their sense of self (Akerlof and Kranton, 2000). This framework helps explain how policy changes that challenge traditional gender roles could affect violence even without changing relative economic resources between spouses: either reducing it by successfully shifting identities and associated prescriptions, or provoking backlash and spurring an IPV increase.

Akerlof and Kranton (2000) define identity as one's sense of self and belonging to social categories, which directly enters utility functions and influences economic decisions. Their model yields three implications particularly relevant for understanding IPV prevention efforts. First, identity can explain seemingly detrimental behaviors that bolster a diminished self-image. For instance, men may use violence to reaffirm masculine identity when it feels threatened. Second, identity creates externalities: one person's behavior change may affect others' identities and provoke responses, potentially generating cascading effects throughout a community. For example, a man taking on housework, childcare, or spending leisure time with his wife instead of at bars may threaten other men's (and women's) identities. This externality may then create further externalities if commu-

nity members react negatively (or positively) to these behavior changes. Third, identity implies a way for preferences to be changed: governments or NGOs could deliberately shift social categories and their prescriptions to reduce gender-based violence, though such efforts may face resistance. These insights suggest IPV prevention programs could either successfully transform gender identities and reduce violence, or trigger backlash if they threaten identities without successfully changing underlying prescriptions.

In applying the identity model to our setting, the IPV prevention couples training program attempts to shift three aspects of identity: the social categories individuals choose (e.g., becoming a 'progressive' man or woman), the relative status of these categories (e.g., elevating women's social position), and the prescriptions for ideal behavior (e.g., men sharing housework and decisions, women having autonomy over sexual choices). Successfully changing these elements could reduce identity-based resistance to women's empowerment and men's more egalitarian behavior. For example, men might face less identity loss from taking on traditionally 'women's work' or spending more time at home, while women might face less resistance when exercising sexual and reproductive autonomy, including initiating or refusing sex. If successful, a program that shifts gender identities in this way could create virtuous cycles of behavior change through positive externalities in the community, ultimately reducing IPV.

However, if programs fail to effectively transform these gendered categories and prescriptions, or if the changed behaviors of program participants provoke a negative reaction in others, then the identity loss could provoke widespread backlash from men. When identity-threatening behaviors spread through a community without corresponding shifts in underlying norms, individuals often respond with social sanctions like gossip, scorn, and ostracism to reaffirm traditional identities (Akerlof and Kranton, 2000). These sanctions can pressure early adopters of progressive behaviors to revert to traditional behaviors and may trigger increased violence as men attempt to reassert threatened masculinity. The identity model thus explains how partial or unsuccessful attempts to shift gender norms could generate widespread negative externalities, including increased IPV.

2.4 Experimental Design and Sampling

The intervention recruited couples through village savings and loan associations (VSLAs) in Rwanda's Eastern Province, working with existing VSLA members supported by CARE International.[8] Each VSLA was associated with a village near its members' homes.[9]

To measure both direct and spillover effects, we used a two-level randomization design: first assigning villages to treatment or control, then randomizing eligible couples within treated villages. This design created three study groups: treated couples (received training), spillover couples (untreated couples in treated villages), and control couples (in control villages), as illustrated in figure 2.

The sampling frame was chosen as follows. The Government of Rwanda, in partnership with province and district authorities, selected four districts and eight sectors in Eastern Province with high IPV rates and VSLA concentrations. Within these sectors, 98 villages were selected that had both sufficient VSLAs (minimum 3) and geographic separation to minimize contamination risk between study villages. Where villages had more VSLAs, the largest ones were selected (see Appendix B for details).

VSLA members were invited to sensitization sessions about a group training on "improving family relationships." Interested couples participated in public lotteries within their VSLAs to determine program eligibility. These lotteries were conducted identically across all villages, ensuring respondents remained blind to their treatment status during recruitment and baseline.

To apply for the couples training, lottery participants had to meet the following eligibility criteria:

1. Married or cohabiting for at least 12 months.

2. At least one partner being a VSLA member for at least three months.

3. Both partners age 18 or over.

4. Both partners willing to attend all 22 weekly sessions.

[8] VSLAs were chosen as an entry point for this program to address financially-controlling behaviors by some husbands of female VSLA members (Stern and Nyiratunga, 2017).

[9] Villages are located below the administrative unit of cells, sectors, districts, and provinces.

MIGEPROF and the research team set the treatment sample size at 640 couples based on budget constraints. The research team first conducted a desk-randomization to assign villages to treatment or control groups using Stata. Then in all study villages, public VSLA lotteries were used to rank eligible couples from 1 to 30 (the maximum VSLA size). The first 60% of couples from each VSLA lottery were assigned to treatment, stratified by sector and number of lottery applicants. The implementing NGO then organized the 640 treated couples into 32 training groups based on location, combining members from nearby VSLAs.

2.5 Representative Baseline Community Survey

To assess selection into the program, we randomly sampled and interviewed 686 couples at baseline from a complete household listing of study villages. VSLA couples were similar to randomly selected non-VSLA couples on most characteristics, though slightly younger, more educated, less economically vulnerable, and more progressive on gender issues (see table A2 for a comparison on key characteristics).

3 Data and Estimation

3.1 Data Collection

Data collection included two rounds of panel surveys. We conducted baseline surveys from November 2017 to February 2018, before the intervention (March–August 2018), and endline surveys six to eight months after completion (February–March 2019).[10] Of the 2,042 baseline study couples, 98.3% of women and 97.5% of men were re-interviewed at endline, with no differential attrition across treatment arms.[11]

Women's and men's surveys covered household demographics, socio-economic characteristics, time use, economic activity, decision-making, well-being, gender attitudes and norms, experience of IPV, and likelihood of perpetrating violence in hypothetical scenarios (men only). Surveys were administered by enumerators of the same sex as the respondent in private settings (empty classrooms, village offices, or churches), with strict privacy protocols requiring interviews to pause

[10] A World Bank qualitative review of implementation was conducted in January 2020.

[11] Village leaders from each 5-person village leadership committee were also interviewed about gender attitudes, infrastructure, and program awareness.

if others came within earshot or eyesight. Informed consent was obtained before both baseline and endline surveys.

3.2 Measuring Violence

We measured IPV using the World Health Organization (WHO) Violence Against Women instrument, following WHO ethical protocols for IPV research (WHO, 2016). The study's ethical protocols are summarized in a structured ethics appendix (Appendix C). This standardized instrument asks multiple behaviorally specific questions about a range of abusive acts, a technique shown to maximize disclosure (Ellsberg et al., 2001). Following the literature, we focused on three types of IPV: emotional (2 questions), physical (7 questions), and sexual violence (3 questions). Women reported how often their current husband perpetrated each act over the past six months at endline (to avoid overlap with program completion) or 12 months at baseline, with responses of 'never', 'once', 'a few times', 'often', or 'skip/refuse to answer'.[12]

Following our pre-analysis plan (Alik-Lagrange et al., 2017), we created binary indicators for each type of violence.[13] These indicator variables were coded as one if a woman reported experiencing any act within that category. When a woman refused to answer some questions but reported violence on others within the same category, the indicator was coded as one. If she refused one question and reported no violence or refused all others, the indicator was coded as missing. Results are robust to coding all refusals as either "yes" or "no" (see table A20 in the appendix).

To minimize reporting bias and increase privacy, we used audio computer-assisted self-interviewing (ACASI) when asking the most sensitive survey questions, including those on IPV, men's likelihood of violence in hypothetical scenarios, alcohol consumption, and gender attitudes. With ACASI, respondents listen to pre-recorded questions through headphones and answer on a tablet touchscreen. Enumerators demonstrate the technology, then stand a short distance away to maintain privacy. This approach is designed to minimize reporting bias by increasing privacy and reducing the embarrassment respondents may feel about disclosing a sensitive answer face-to-face with an enumerator

[12] Examples of questions included: "In the past six months, how often has your husband said or done something to humiliate you in front of others", "punched you", or "physically forced you to have sexual intercourse with him when you did not want to."

[13] Due to a coding error, one emotional and one physical violence question was only asked of a third of endline respondents and was dropped from the emotional and physical violence endline dummies. See Appendix D for deviations from the pre-analysis plan.

(Cullen, 2023).[14]

We also measured impacts on secondary outcomes suggestive of mechanisms. For conceptually linked outcomes measured with multiple questions, we constructed summary indices following Anderson (2008) and our pre-analysis plan, to minimize risks from multiple hypothesis testing. For other related outcomes, where more appropriate, we summed the number of agreed statements within a category (e.g., opinion about situations where wife-beating is acceptable) or reported individual question scores.

3.3 Sample Characteristics

Table A3 shows baseline balance across treatment groups, with no normalized differences exceeding 0.25 standard deviations, indicating good balance (Imbens and Rubin, 2015). Men and women averaged 40 and 36 years old with 4.7 and 4.5 years of education respectively. Half the sample was legally married, with others cohabiting. Couples had an average of 3.7 children. Baseline IPV rates were high: in the preceding 12 months, 49% of women reported emotional violence, 40% physical violence, and 42% sexual violence.

3.4 Program Take-Up and Compliance to Treatment Assignment

Program take-up and attendance were high: 95% of couples offered a place attended the training, completing an average of 21 of 22 sessions.

As treatment assignment was based on VSLA membership, some couples belonged to VSLAs in villages different from their official residence. This is common in Rwanda's densely populated context, where VSLA membership is not restricted by village boundaries. We accounted for this VSLA-village structure by implementing geographic buffers between study villages, minimizing the extent of this partial 'compliance'.[15] Our results are robust to various specifications accounting for this geographic distribution (tables A4, A5, A6).

[14] Cullen (2023) found ACASI elicited marginally higher reports of sexual violence than face-to-face methods in this context, though lower than list experiments, which were not used here due to power and reliability concerns (Chuang et al., 2021; Cullen, 2023).

[15] Over 70% of control couples had no treated couples within 600 meters of their dwelling (a significant distance given the region's terrain), rising to 90% for a 200-meter radius (the radial densities are shown in figure A1).

3.5 Econometric Specification

We exploit the two-level randomization to identify direct and spillover effects, estimating intention-to-treat (ITT) effects using an analysis of covariance (ANCOVA) model:

$$Y_{i,v}^{Post} = \theta Y_{i,v}^{Pre} + \beta T_{i,v} + \delta S_{i,v} + X_{i,v}'\gamma + \epsilon_{i,v}$$

where $Y_{i,v}^{Pre}$ and $Y_{i,v}^{Post}$ are baseline and endline outcomes for individual i in village v, $T_{i,v}$ and $S_{i,v}$ are indicators for treatment and spillover assignment, and $X_{i,v}$ is a vector of baseline or time-invariant controls and strata fixed effects.

The coefficients β and δ capture treatment and spillover ITT effects respectively. To test whether β and δ are equal, we conduct Wald tests and report the p-values. We cluster standard errors at the village level and adjust for multiple hypothesis testing across primary outcomes using the Benjamini and Hochberg (1995) procedure (Anderson, 2008). Results are robust to including additional controls, including those selected by post-double selection LASSO. For outcomes only measured at endline, we estimate effects using a single difference model.

4 Results

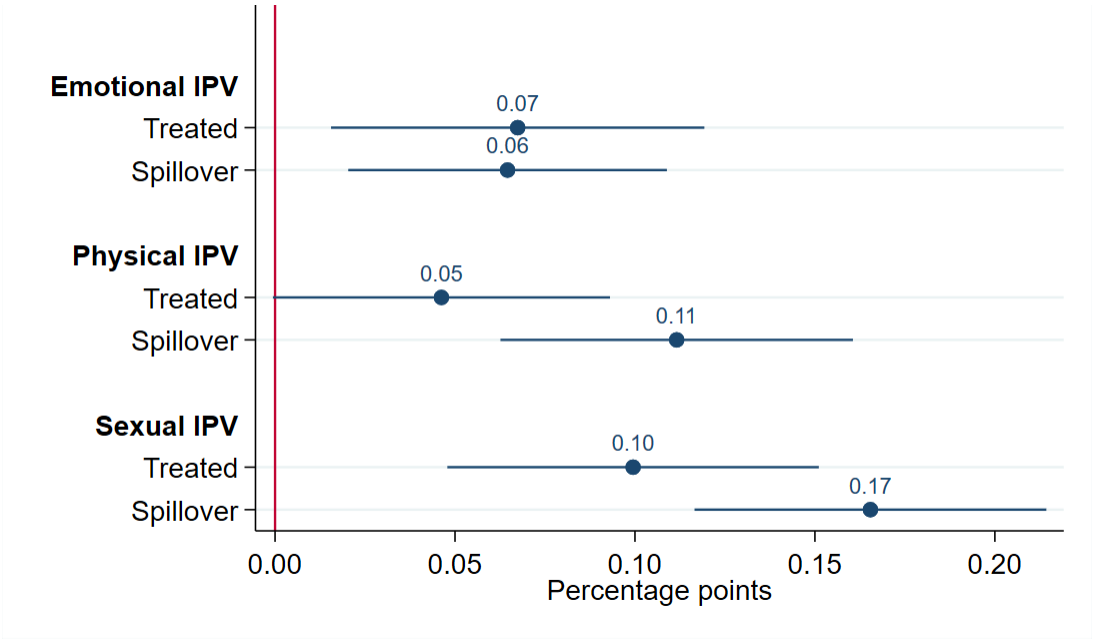
4.1 Primary Outcomes: Large, Unintended Increases in IPV

We find that six months after the program ended, the intervention substantially increased women's self-reported experience of IPV (table 1 and figure 1). For treated couples, the intervention caused 7, 5, and 10 percentage points increases in emotional, physical, and sexual violence, respectively, compared to the control group. For spillover couples, the intervention caused a seven percentage points increase in emotional violence, and significantly larger increases in physical and sexual violence, of 11 and 17 percentage points respectively. Relative to the control group, these impacts represent large increases of 15% and 33% in physical violence and 24% and 40% in sexual violence for treated and spillover couples respectively.

Related outcomes also worsened (table A7). Women reported more controlling behaviors from husbands (for spillover women), more injuries from sexual violence, and more perceived IPV incidents in their village. Enumerators also observed more visible physical injuries among spillover

women. Violence against men increased as well: treated men reported eight percentage points more emotional abuse from wives, while spillover men reported seven percentage points more physical violence.

Figure 1: Treatment effects on IPV prevalence



Note: IPV = intimate partner violence. This figure shows treatment and spillover effects on the three IPV outcomes. Data are expressed as percentage points treatment effects relative to the control group, with 95 percent confidence intervals. Full statistical results for the treatment effects are presented in table 1. The outcome variables are bolded, with the coefficients for treated and spillover below. The dependent variable is a dummy aggregating whether a wife reported experiencing any of each of the forms of emotional, physical or sexual violence, controlling for the baseline value of the dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.

4.2 Secondary Outcomes

The intervention also negatively affected several relationship and well-being outcomes (table 2). While an index of overall relationship quality was unchanged, both treated and spillover women reported more marital disagreements escalating to violence and more drunken quarrels. Consistent with this, both groups of women reported more excessive drinking by husbands, and men's self-reports confirm increased alcohol consumption among spillover men (table 3).

Both groups of women also reported husbands more frequently spending household money on themselves, indicating deteriorating economic cooperation. Overall, these effects tend to be significantly worse for spillover couples, who also experienced additional adverse impacts: while treated women gained decision-making power, spillover women lost it (table 4). Spillover couples also reported increased depression (both spouses) and husband infidelity. These negative effects are not found among treated couples (table 2).

Economic impacts were minimal. Couples reported slightly more household assets (table A8), and treated women's savings increased by a small but statistically significant amount (equivalent to 6 USD), suggesting that treated women, but not their husbands, may have saved some of the travel stipend. There were no effects on women's asset share or economic participation.[16]

5 Mechanisms

In this section, we discuss the evidence on potential mechanisms that may be driving the increase in IPV. We then discuss alternative mechanisms, including an in-depth investigation of whether increased reporting of IPV rather than a true increase in IPV might be driving results.

5.1 Women's Progressive Aspirations and Attitudes, and Men's Challenged Identity

Suggestive evidence indicates that increased IPV stemmed from male backlash against threatened gender identities posed by the program's messages about gender equality. This backlash appears to have been community-wide, as indicated by large negative spillover effects, and to have been partially mitigated in treated couples, with these men adopting some new progressive attitudes and behaviors encouraged in the training.

5.1.1 Qualitative Reports

To understand the program's unintended negative effects, we commissioned an independent ex-post review. The qualitative data discussed here were gathered a year after the program ended. It illus-

[16]The program aimed to increase cooperation over household budgets and women's decision-making input in income use. While about a third of treated households reported saving some of the $2 per person weekly stipend, this amount was smaller than standard cash transfers known to have minimal economic impacts in a similar setting (McIntosh and Zeitlin, 2024).

trates how the program challenging traditional gender roles may have triggered violence, broadly consistent with our quantitative findings.

Men's Resistance to Losing Power

According to facilitators interviewed, men strongly resisted challenges to their power while " women were excited that men's power was being challenged [and] felt like they were finally learning the reality of the situation." The curriculum began by examining power dynamics and distinguishing between biological sex and cultural gender prescriptions. While this initially confused participants — "most couples thought gender and sex were the same thing" — understanding this distinction "changed the dynamics of the class." However, a facilitator noted, "It was difficult for men to accept that gender is not biological, that it's culture and patriarchy that give more power to men." Men particularly resisted sharing domestic responsibilities as violating 'natural' gender roles, arguing, "In our culture, it is the man who pays the [bride price]. Are you now saying that women can also pay the [bride price]?"

Qualitative reports also suggest that some trained men initially adopted new behaviors encouraged by the program, such as sharing domestic and care work or reducing drinking at bars. However, facilitators reported these men faced social sanctions: they were shamed as "not real men," accused of doing "women's work," and rumored to be bewitched by their wives. According to the facilitators, this ostracism deterred other men from changing and led early adopters to revert to traditional behaviors. As one participant explained: "If a man is very serious about the program, he can say, 'It's fine if my wife bewitches me. I don't care, as long as I don't die.' But most of the men reverted back to their old behavior and are doing rough things to their families."

Dissemination in the Community

Program messages spread widely through both formal and informal channels. Village leaders encouraged participating couples to share their experiences at regular community-wide meetings and with friends and religious communities. Local government officials were supportive, encouraged by mandate to address gender-based violence in their *imihigo* government performance contracts. Leaders and facilitators observed that couples shared information during various community events:

informal gatherings at bars, churches, or markets; VSLA meetings; and regular pre-existing community meetings like parenting meetings and compulsory monthly whole-of-village *umuganda* community service gatherings.[17]

These public discussions often provoked resistance, particularly around women's sexual autonomy. Community members sometimes openly challenged such messages during whole-of-village meetings. As one facilitator noted, men "often think that this thing of gender has just come to bring problems to the man and to take away his rights." Another warned about introducing program ideas without adequate explanation: "When you start talking about [gender-based violence] in a local community and you don't adequately explain the issue, they'll think that you're just stripping us of our power and property. Local communities can [exact] revenge if they are introduced to these ideas without fully understanding them."

The program's visibility was amplified by Rwanda's high population density. Study villages (typically with 150–200 households) averaged 27 participating couples, with some having up to 41, comprising up to one-third of a village. Most spillover couples (67% of men, 69% of women) knew program participants personally, and nearly half discussed the curriculum with them. Facilitators noted that many women in the community, observing changes in participating peers, expressed interest in joining.

5.1.2 Quantitative Results

The qualitative reports — suggestive of a mechanism of backlash against threatened identity, men's loss of power, and shaming of those who adopt new behaviors — are supported by the quantitative data. As predicted by the identity model (section 2.3), men may respond to threats to traditional masculinity by acting out against women and shaming men who adopt progressive behaviors, leading to increased violence, both in the broader community and program participants. To offset the disutility caused by changes that threaten traditional male identity, men may seek to reaffirm that identity through violence and by publicly shaming those adopting new behaviors. This could in turn lead to reversion to old behaviors, including increased IPV. Our data are consistent with two complementary mechanisms: first, the village-wide diffusion of progressive attitudes and be-

[17] In these settings, couples typically shared how behavior changes like joint financial decisions improved their household finances and marriages, rather than discussing theoretical concepts like IPV definitions.

haviors that challenge traditional gender prescriptions, and second, increased friction in couples as spouses diverge in their attitudes and aspirations for power within the household. Both mechanisms threaten male identity and may trigger violent responses.

We first examine evidence of backlash among spillover couples before turning to treated couples. This ordering reflects both the importance of community norms in driving backlash, and the more straight-forward interpretation of spillover effects (for example, treated men show mixed responses including increased progressive attitudes).

Backlash among Spillover Couples

First, we find that spillover men report more negative attitudes than control men (panel B of table 5): they justify IPV in 0.3 more situations and hold 0.1 more gender-unequal attitudes (marginally significant). Further, they report being more likely to perpetrate violence against their wife in 0.95 additional scenarios (against a control mean of 3.5). These worsened attitudes and hypothetical violent behavior align with spillover women's reports of increased violence.

Spillover men also exhibit greater desire for household control (table 4). They report that in an ideal world, they would want more input in decision-making about household expenses (3 percentage points above control). They are also more likely to claim to be the sole decision-maker across three of four domains, including household income use and major expenses (figure 3). This shift toward sole decision-making is striking given the low baseline rates in control villages (3–7 percent); men who are claiming to be sole decision-makers are thus going against the prevailing norm of joint decision-making. These changes in attitudes and decision-making suggest spillover men are asserting more household power and violence, as a reaction to a threat to their identity.

To further corroborate the threatened identity and backlash mechanism, we examine evidence of intra-household friction between increasingly progressive wives and resistant husbands. Indeed, spillover women report significantly more progressive gender attitudes and reduced IPV acceptance, while their husbands become more conservative (panels A and B of table 5). Specifically, spillover women justify IPV in 0.2 fewer situations and hold 0.1 fewer gender-unequal attitudes than control women. This husband-wife divergence in attitudes is significantly larger in spillover couples: the gap in IPV acceptability between spouses is 0.6 reasons greater, compared to a control mean of 1.19

(panel C of table 5, figure 4). Moreover, we find a five percentage points increase in the share of couples where wives became more progressive while husbands became more conservative between baseline and endline. Spillover women also report a nine percentage points higher desire for more input in household financial decisions (table 4, figure 5), directly conflicting with their husbands' increased desire for control. These findings align with evidence from India showing that larger differences in couples' attitudes and men's social image concerns drive backlash behavior (Cullen et al., 2024). The negative spillover effects we document also accord with those identified by Chatterji et al. (2020), who found increases in IPV among randomly selected community members after a similar program in Rwanda.

If the IPV backlash stemmed from threatened male identity, we would expect the largest effects among couples not already experiencing violence at baseline. Men who were already perpetrating IPV should face little additional motivation to use IPV to assuage threatened identity, while men not perpetrating IPV at baseline may have greater social returns from using IPV to assuage an existing (or acquired) reputation as being more progressive. Indeed, negative treatment effects are greatest for spillover couples where women reported no IPV at baseline (table A9). For women who did report baseline IPV, the intervention also increased the severity of the type of violence they experienced, also consistent with the mechanism of men assuaging threatened identity with worsened IPV.[18]

Overall, the evidence strongly supports a backlash mechanism among spillover couples. Quantitative results, corroborated by anecdotal reports, indicate increased friction as women's and men's attitudes diverged and aspirations clashed. Moreover, the diffusion of progressive messages by trained couples and through community events threatened spillover men's identity, leading them to reassert dominance through violence.

Evidence on Village-Level Dissemination Effects

Not all couples belonging to a VSLA live in the same village as their VSLA's location. This geographic variation enables us to examine the role of the village as a channel for program message dissemination causing spillovers, by studying couples who reside in treated villages but belong to

[18]For example, women experiencing only emotional IPV at baseline escalated to physical and/or sexual IPV at endline. Similarly, women experiencing physical but not sexual IPV at baseline reported increased sexual IPV at endline.

VSLAs in control villages. Since selection into village of residence may be endogenous, we leverage the random assignment of villages to treatment or control. By using village treatment assignment as an instrument for actual village treatment status, we can estimate treatment-on-the-treated (ToT) effects for spillover couples whose VSLA was located in a different village from their residence.

The two-stage least squares analysis reveals substantially larger negative spillover effects for spillover couples who actually reside in treated villages: emotional IPV increases by 12 percentage points, physical IPV by 19 percentage points, and sexual IPV by 27 percentage points (table A6), all much higher than the ITT spillover effects. The much larger ToT spillover effects suggest that geographic proximity to program implementation, rather than VSLA membership alone, drives negative spillovers, highlighting the importance of village-level transmission channels for program messaging and norms.

Backlash among Treatment Couples

Mechanisms for treated couples differ slightly from those of spillover couples, though we still find evidence consistent with backlash effects, potentially alongside other mechanisms. Unlike spillover men, and in accordance with the prevention program's theory of change, treated men hold fewer conservative gender attitudes and report being less likely to perpetrate violence against their wife than control men (panel B of table 5). The program similarly causes their wives to become much more progressive on gender issues. Women in the treatment arm report the greatest number of progressive gender attitudes and least acceptance of IPV compared to control or spillover women. Treated women support 0.4 and 0.6 fewer gender unequal attitudes and justifications for IPV, respectively, than control women. This compares to 0.1 and 0.23 fewer conservative attitudes, respectively, for spillover women.

While both treated men and women become more progressive in their gender attitudes, women progress significantly more: the program increases the number of additional progressive attitudes that treated women hold compared to their husbands regarding both gender roles (0.2 attitudes) and IPV acceptability (0.4 reasons). This divergence suggests that treated couples, like spillover couples, experienced increased friction as women's attitudes shifted more than their husbands' (panel C of table 5, and figure 4). This attitude divergence, similar to spillover couples, may have triggered backlash violence from men feeling threatened. Alternatively, treated men's behaviors

may diverge from their stated attitudes and intentions because they still lack the skills to enact change, despite wanting to.[19]

One exception to the general pattern of improved attitudes concerns sexual autonomy: both treated men and women believe wives are justified in refusing sex in fewer situations than the control group. This unique domain where attitudes worsened across all groups aligns with qualitative evidence of strong public resistance to messages about women's sexual rights.[20] While this consistent negative shift may not have increased friction within couples, it could have contributed to backlash through hearing messages on this topic in the training and community fora.

The program also increased treated women's aspirations for agency, similar to spillover women. Treated women wanted more input in household expense and childbearing decisions (six and five percentage points respectively; table 4 and figure 5). While treated women's husbands, unlike spillover men, did not seek more decision-making power, even this smaller increase in treated women's aspirations for agency points to a potential shift in dynamics within the household power structures that may have been enough to trigger backlash.

We now turn to changes in treated men's behaviors. Program effects on treated men differ markedly from spillover men, with some progressive changes. However, these new attitudes and behaviors adopted by treated men may have exposed them to community shaming, potentially pressuring them to reassert traditional identities through violence.

Table 3 reveals several progressive behavioral changes among treated men. Time-use diaries show both treated men and their wives increased time spent on domestic and care work, traditionally considered women's responsibility. This was aligned with the program's emphasis on positive parenting and active engagement in child development, and represented a behavior shift potentially visible to other community members. The program also enhanced treated women's household decision-making power, as reported by both spouses (table 4).

Marriage formalization provides further evidence of publicly visible behavior change: among couples not formally married at baseline (50% of the sample), treatment increased formalization by 24 percentage points, with no effect for spillover couples. While mass weddings are common

[19] Another possibility is social desirability bias: men may have learned the 'right' answers during training. If so, the true attitude divergence between spouses could be even larger, assuming women's responses face less such bias.

[20] Facilitators identified sexual relations as particularly controversial: "there's a commonly held belief that a man should initiate sex. Men were resistant to the idea that women can initiate sex and that men don't just have a right to sex whenever they want it".

in Rwanda, where formal marriage protects women's rights,[21] program effects were concentrated during the training period, with 36% occurring in the final month.[22]

These progressive and potentially publicly observable behavior changes, absent among spillover couples, may have paradoxically increased treated men's IPV through community backlash. Men who adopted new behaviors such as doing more domestic and care work, or by getting married, may have faced social sanctions from others in their communities, including spillover men whose attitudes had worsened. This shaming may have driven treated men to reassert traditional masculinity through violence. Consistent with this, treated men's increased alcohol consumption aligns with qualitative reports of drinking as a masculinity-asserting behavior (table 3). Overall, treated couples show two pathways to increased IPV: directly, through friction from diverging attitudes within couples, and indirectly, through pressure from community shaming of progressive behaviors—consistent with evidence linking men's social image concerns to backlash against gender empowerment programs.

5.2 Alternative Explanations

5.2.1 Distinguishing True Increases in IPV from Changes in Reporting

In this section, we present findings from a systematic analysis to assess whether our results reflect actual increases in violence rather than changes in reporting behavior, and find evidence consistent with a true increase in IPV. This novel, comprehensive approach to addressing measurement concerns in IPV reporting is important given potential correlation between treatment and IPV reporting.

Treatment could affect reporting bias through three main channels: (1) strategic reporting (to access future programs and stipends); (2) expanded understanding of IPV question terms; and (3) reduced stigma about disclosure or empowerment to report. Each could bias our treatment effects if correlated with treatment status. We examine evidence for each below.

[21] Women in informal marriages have weaker property and inheritance rights (Kaiser-Hughes et al., 2016). This is widely recognized: 84% of Rwandans know informal marriages offer women fewer legal protections, and only 43% support equal property rights for informally married women versus 94% for formally married women (USAID, 2016).

[22] Facilitators reported organizing joint ceremonies and noted marriage's importance: "If a woman is legally married, she's more confident in challenging her husband... It imposes some limit on men when marriages are formal/legal."

Strategic Reporting

One concern is that couples might over-report IPV to increase chances of accessing future program benefits, given the attendance stipend. However, if couples were reporting strategically to appear as favorable program candidates, we would expect them to minimize reports of socially undesirable behavior. Instead, we find the opposite: both treated and spillover couples report increased alcohol consumption by husbands across multiple measures (table 3).

Strategic reporting would also suggest couples might claim greater poverty to justify need for program stipends. Instead, we find either no economic effects or slight improvements: both treated and spillover couples report more assets, and treated women report increased savings (table A8).

Moreover, strategic coordination would likely produce aligned responses between spouses on key program themes. Yet we observe significant divergence in gender attitudes between husbands and wives. If there were shared community understanding about reporting to gain program access, we would expect similar responses across all groups. Instead, we find different directions of impacts between spouses and between treated and spillover couples, with mechanisms and magnitudes differing across groups—inconsistent with coordinated strategic reporting. For example, we find diverging effects on gender attitudes across groups and genders, yet worsened acceptance of the justification for women refusing sex in all groups.

Finally, strategic reporting implies couples believed their responses affected program enrollment. However, the transparent public lottery process, confirmed by village leaders' recollections and baseline balance in key variables, makes this unlikely.

Expanded Understanding of IPV Questions

Second, the intervention may have expanded women's understanding of IPV survey questions. If so, terms like "slapping," "choking", "force," or "degrading" may have taken on different meanings for women in treated versus control groups. However, several patterns suggest expanded understanding does not drive our results. First, we find the largest effects among spillover women who had only indirect exposure to IPV discussions, rather than treated women who had direct program exposure.[23]

[23]Qualitative reports indicate that in community meetings, couples mainly shared personal testimonials about changes in their relationships rather than discussing IPV definitions and concepts.

Second, effects are similar across questions using subjective versus objective terms, though expanded understanding should primarily affect subjective interpretations (table A10). We observe increases across the spectrum of individual behaviorally specific questions, encompassing both subjective and objective questions. Finally, while both treated and spillover women report more frequent and additional types of violence than the control group, these increases are again largest for spillover women (table A11). This pattern—where those with the most comprehensive IPV knowledge (treated women) report less violence than those with indirect exposure—is inconsistent with expanded understanding driving our results.

Reduced Stigma or More Empowered to Report

Finally, treatment might have made women more comfortable reporting IPV through reduced stigma or a sense of increased empowerment.[24] However, several patterns contradict this explanation. Tables A12 and A13 show that while treated women are marginally less likely to refuse to answer physical IPV questions, spillover women are marginally more likely to refuse, with no differences for other IPV types. Further, while treated women are seven percentage points more likely to visit safe spaces or speak about violence they experienced, they are less likely to report witnessing IPV to local leaders. Similar safe space usage by treated men suggests these were seen as general marital support rather than stigmatized services. Overall, there is no evidence that spillover women - who report the highest IPV rates - show signs of reduced reporting stigma or increased empowerment in reporting.

Additional Evidence on Truth or Reporting

Several additional patterns support our interpretation of true violence increases rather than reporting changes. First, women's IPV reports correlate positively with observable indicators: distress during the interview and referral to a counselor, choosing to speak to the counselor, and showing visible signs of injury observed by the enumerator (correlation coefficients are in table A14). While there were few women with these characteristics, leaving us underpowered to detect treatment effects, the positive correlations suggest reported IPV reflects actual violence rather than just

[24]This could reflect a positive impact if underreporting prevents victims from accessing services.

reporting differences. Moreover, the treatment effects show enumerators identified more visible injuries among spillover women (table A7), further corroborating the IPV increases.

Effects also vary predictably with couples' time together. At baseline, 230 men reported working away from home for four or more consecutive nights in the previous three months. As most IPV requires physical presence, we would expect larger effects when husbands are home more frequently. Indeed, physical and sexual IPV increases are substantially larger among spillover couples where husbands did not migrate for work (table A15).[25] These wives also report more types and higher frequency of sexual IPV—both patterns unlikely to arise from reporting bias alone.

Finally, the treatment effects reveal a coherent pattern across multiple related measures: increased sexual IPV, sexual violence injuries, perceived community IPV prevalence, worsened attitudes about wives' right to refuse sex across all groups, and more sex-related arguments in couples. This consistency across diverse related indicators, combined with other objectively negative outcomes for which misreporting is unlikely, like depression, infidelity, and alcohol use, further supports our interpretation of true violence increases rather than reporting changes.

5.2.2 Changing Descriptive Norms

An alternative to the backlash mechanism is that this large-scale government program may have inadvertently 'loosened' social constraints around IPV by updating perceptions of its prevalence. While baseline beliefs significantly underestimated community IPV rates (table A2), the intervention increased all groups' perceptions of its pervasiveness (table A16). This updating about IPV's commonality might have reduced perceived social sanctions and enabled more violence through increased peer acceptance (Bursztyn et al., 2020; Mulla et al., 2019; Tankard and Paluck, 2016).

However, if belief updating was the main driver of violence increases, effects should be largest for treated couples who updated their beliefs most substantially. Instead, we find the opposite pattern: spillover couples show larger IPV increases despite smaller updates in beliefs. The absence of heterogeneous treatment effects by baseline norm beliefs further suggests this mechanism is not the primary one at play.

[25] This pattern appears only for spillover couples whose routines were not affected by training attendance, perhaps explaining why we do not see differential effects for treated couples.

5.2.3 Resurfaced Trauma

The program's discussion of sensitive topics might have triggered past traumas, increasing stress and violence. We examine this possibility through heterogeneous treatment effects based on baseline variables that proxy past traumas measured by husbands' negative experiences during the 1994 genocide and childhood exposure to IPV (table A17). We find no evidence that pre-existing trauma predicts larger treatment effects, though this exploratory analysis may be underpowered to detect such relationships.

5.2.4 Violence as Response to Women's Economic Gains

Economic theory and previous research suggests women's increased financial resources can trigger IPV through either instrumental extraction or threatened-identity–based retaliation (Angelucci, 2008; Bobonis et al., 2013; Hidrobo and Fernald, 2013). However, while treated women's savings increased modestly (equivalent to USD 6), spillover women experienced the largest violence increases despite no economic gains, suggesting this is not a primary mechanism in our context.

5.2.5 Implementation Factors

Suggestive positive effects from other IPV prevention programs, leaving the aforementioned methodological questions aside, may indicate that small differences in context, implementer, or program delivery could substantially affect impact direction (Doyle et al., 2018; Dunkle et al., 2020; Sharma et al., 2020). The version tested here differed in several ways from the evaluation of a similar earlier program that found reductions in treated couples' IPV and suggestive increases in community IPV (Dunkle et al., 2020).[26] The program evaluated in this paper had larger groups (19 couples on average), shorter facilitator training, and potentially additional community message diffusion through village meetings not sanctioned by the program. These features may have hindered the program's ability to contain backlash risks. Regardless, our results show that scaling up IPV prevention through government, even when implemented by experienced gender-focused NGOs, can pose significant risks for harm.

[26] Another noteworthy difference between the two studies is that not all components of the program had been fully delivered by the time of their evaluation in Dunkle et al. (2020), whereas in this study all interventions had completed delivery by endline.

6 Other Robustness Checks

Treatment effects on primary IPV outcomes are stable across multiple specifications (tables A18, A19). Effect magnitudes and precision are consistent whether using single difference models, post-double selection LASSO controls (Belloni et al., 2014), probit estimation, or estimating treatment-on-the-treated, instrumenting treatment with assignment. Additionally, results are unchanged when coding all IPV question refusals as either "yes" or "no" (table A20), indicating findings are not driven by differential response rates.

7 Conclusion

This paper studied the effect of a 22-week couples discussion program designed to prevent IPV in rural Rwanda. We found large, unintended increases in IPV for both treated couples and spillover couples in treated villages. The intervention increased women's reported experience of emotional, physical, and sexual IPV, with effects largest for control couples in treated communities. Systematic analysis suggests these effects reflect true increases in violence rather than changes in reporting behavior. Negative effects spanned a range of outcomes including increased alcohol consumption, arguments, violence against men, visible injuries, infidelity, and depression.

Our findings demonstrate how threats to traditional identities and social norms can drive behavior. While some research finds evidence of backlash against women's economic empowerment, we demonstrate that perceived identity threats alone—without changes in relative resources—can trigger substantial retaliatory behavioral responses. Evidence is consistent with the program threatening traditional gender identities, prompting widespread backlash through two main channels: increased friction as couples' attitudes diverged, and community resistance to progressive messaging, which may have triggered social sanctions on men adopting new behaviors.

Our study makes two broader contributions. Methodologically, by measuring spillovers, we show how standard impact evaluations focused only on direct beneficiaries might miss crucial community dynamics. Our results also suggest that similar programs challenging sensitive traditional social structures, such as gender, caste, or ethnicity, and where higher-status individuals feel they are 'losing out,' may face similar resistance if they threaten established identities.

The magnitude of negative impacts from an IPV prevention program that followed best practices and careful contextual adaptation has critical policy implications. As similar programs scale across Sub-Saharan Africa and beyond, our results suggest the need for careful monitoring and better understanding of underlying mechanisms to minimize backlash risks. Recent evidence suggests emphasizing broader community benefits of a program, rather than women's rights as the primary objective, might reduce perceived zero-sum losses and men's subsequent retaliation (Cullen et al., 2024). Future research could seek to understand ways to mitigate backlash risk and better understand community transmission mechanisms. While our results serve as a cautionary tale about unintended negative indirect effects, they also suggest that if such diffusion mechanisms could be harnessed positively, targeted programs might achieve broader community impact cost-effectively (e.g. Andrew et al., 2022).

Alternative approaches may also prove effective. In addition to continuing to refine and carefully test small-scale programs seeking to transform gendered power relations and norms, researchers should investigate potentially less risky alternatives. Adolescent-focused programs may face lower resistance, as younger individuals have more flexible gender attitudes and boys may perceive less threat from future wives' empowerment (Andrew et al., 2022; Dhar et al., 2022). Pre-marriage interventions might also improve outcomes through improved partner quality and matching (Bandiera et al., 2020; Shah et al., 2023). Several non-gender-focused approaches may also reduce IPV without threatening, identity such as improving education and employment opportunities, eased divorce laws, or rural electrification expanding access to phones and media (Buller et al., 2018; Sanin, 2021). Understanding which approaches most effectively reduce IPV without triggering identity-based backlash remains a crucial challenge for future research and policy.

References

Abbott, Pamela and Dixon Malunda, "The Promise and the Reality: Women's Rights in Rwanda," *African Journal of International and Comparative Law*, 2016, *24* (4), 561–81.

Abramsky, Tanya, Karen Devries, Ligia Kiss, Janet Nakuti, Nambusi Kyegombe, Elizabeth Starmann, Bonnie Cundill, Leilani Francisco, Dan Kaye, and Tina Musuya, "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 Medicine*, 2014, *12* (1), 122.

Adams, Abi, Kristiina Huttunen, Emily Nix, and Ning Zhang, "The dynamics of abusive relationships," *The Quarterly Journal of Economics*, 2024, *139* (4), 2135–2180.

Aizer, Anna, "Poverty, Violence, and Health the Impact of Domestic Violence during Pregnancy on Newborn Health," *Journal of Human resources*, 2011, *46* (3), 518–538.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, 2000, *115* (3), 715–53.

Alik-Lagrange, Arthur, Claire Cullen, Muthoni Ngatia, and Julia Vaillant, "Preventing Intimate Partner Violence: Impact Evaluation of MIGEPROF Couples Training for IPV Prevention in Eastern Rwanda'. AEA RCT Registry," 2017.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–95.

Anderson, Siwan and Garance Genicot, "Suicide and Property Rights in India," *Journal of Development Economics*, 2015, *114*, 64–78.

Andrew, Alison, Sonya Krutikova, Gabriela Smarrelli, and Hemlata Verma, "Gender Norms, Violence and Adolescent Girls' Trajectories: Evidence from a Field Experiment in India," Technical Report, IFS Working Papers 2022.

Angelucci, Manuela, "Love on the Rocks: Domestic Violence and Alcohol Abuse in Rural Mexico," *Journal of Economic Analysis Policy Contributions*, 2008, *8.*

Asiedu, Edward, Dean Karlan, Monica Lambon-Quayefio, and Christopher Udry, "A Call for Structured Ethics Appendices in Social Science Papers," *Proceedings of the National Academy of Sciences*, 2021, *118* (29), e2024570118.

Asiedu, Elizabeth, Claire Branstette, Babulal Neepa Gaekwad, and Malokele Nanivazo, "The Effect of Women's Representation in Parliament and the Passing of Gender Sensitive Policies," in "AEA Paper Session" 2018. Issues on African Development I, 5 January. Philadelphia.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa," *American Economic Journal: Applied Economics*, 2020, *12* (1), 210–259.

Banerjee, Abhijit, Eliana Ferrara, and Victor Orozco, "Entertainment, Education, and Attitudes toward Domestic Violence," in "AEA Papers and Proceedings," Vol. 109 2019, p. 133–37.

Baranov, Victoria, Lisa Cameron, Diana Contreras Suarez, and Claire Thibout, "Theoretical Underpinnings and Meta-Analysis of the Effects of Cash Transfers on Intimate Partner Violence in Low- and Middle-Income Countries," *Journal of Development Studies*, 2020. Routledge.

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova, "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India," *Science*, 2012, *335.*

Becker, G., *A Treatise on the Family*, Harvard University Press, 2009.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Benabou, R. and J. Tirole, "Laws and Norms.," *National Bureau of Economic Research w17579*, 2011.

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B (Methodological*, 1995, *57* (1), 289–300.

Bhalotra, Sonia, Uma Kambhampati, Samantha Rawlings, and Zahra Siddique, "Intimate partner violence: The influence of job opportunities for men and women," *The World Bank Economic Review*, 2021, *35* (2), 461–479.

Bloch, Francis and Vijayendra Rao, "Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India," *American Economic Review*, 2002, *92* (4), 1029–43.

Bobonis, Gustavo J, Melissa González-Brenes, and Roberto Castro, "Public Transfers and Domestic Violence: The Roles of Private Information and Spousal Control," *American Economic Journal: Economic Policy*, 2013, *5* (1), 179–205.

Boyer, Christopher, Elizabeth Levy Paluck, Jeannie Annan, Tvisha Nevatia, Jasper Cooper, Jackline Namubiru, Lori Heise, and Rachel Lehrer, "Religious Leaders can Motivate Men to Cede Power and Reduce Intimate Partner Violence: Experimental Vvidence from Uganda," *Proceedings of the National Academy of Sciences*, 2022, *119* (31), e2200262119.

Buller, Ana Maria, Amber Peterman, Meghna Ranganathan, Alexandra Bleile, Melissa Hidrobo, and Lori Heise, "A Mixed-Method Review of Cash Transfers and Intimate Partner Violence in Low- and Middle-Income Countries," *The World Bank Research Observer*, 2018, *33* (2), 218–58.

Bulte, Erwin and Robert Lensink, "Women's Empowerment and Domestic Abuse: Experimental Evidence from Vietnam," *European Economic Review*, 2019, *115* (June), 172–91.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American economic review*, 2020, *110* (10), 2997–3029.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka, "The Long-Run Effects of Disruptive Peers," *American Economic Review*, 2018, *108* (11), 3377–3415.

Chatterji, Sangeeta, Erin Stern, Kristin Dunkle, and Lori Heise, "Community Activism as a Strategy to Reduce Intimate Partner Violence (IPV) in Rural Rwanda: Results of a Community Randomised Trial," *Journal of Global Health*, 2020, *10* (1).

Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban, "Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods," *Journal of Development Economics*, 2021, *148*, 102582.

Cools, Sara and Andreas Kotsadam, "Resources and Intimate Partner Violence in Sub-Saharan Africa," *World Development*, 2017, *95* (July), 211–30.

Cullen, Claire, "Method Matters: The Underreporting of Intimate Partner Violence," *The World Bank Economic Review*, 12 2023, *37* (1), 49–73.

_ , Sarthak Joshi, Joseph Vecci, and Julia Talbot-Jones, "Female Empowerment and Male Backlash: Experimental Evidence from India," *Working Paper*, 2024.

Dhar, Diva, Tarun Jain, and Seema Jayachandran, "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," *American Economic Review*, 2022, *112* (3), 899–927.

DHS, "Rwanda Demographic and Health Survey 2014-15'," 2016.

Doyle, Kate, Ruti G. Levtov, Gary Barker, Gautam G. Bastian, Jeffrey B. Bingenheimer, Shamsi Kazimbaya, and Anicet Nzabonimpa, "Gender-Transformative Bandebereho Couples," *Intervention to Promote Male Engagement in Reproductive and Maternal Health and Violence Prevention in Rwanda: Findings from a Randomized Controlled Trial'. PLOS ONE*, 2018, *13* (4), 0192756.

Dunkle, K., E. Stern, and L. Heise, "Impact of the Indashyikriwa Training Curriculum for Couples on Intimate Partner Violence in Rwanda: Cluster Randomized Control Trial," 2020.

Ellsberg, Mary, Lori Heise, Rodolfo Pena, Sonia Agurto, and Anna Winkvist, "Researching Domestic Violence against Women: Methodological and Ethical Considerations," *Studies in Family Planning*, 2001, *32* (1), 1–16.

Erten, Bilge and Pinar Keskin, "For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey," *American Economic Journal: Applied Economics*, 2018, *10* (1), 64–105.

_ and _ , "Trade-offs? The impact of WTO accession on intimate partner violence in Cambodia," *The Review of Economics and Statistics*, 2021, pp. 1–40.

Eswaran, Mukesh and Nisha Malhotra, "Domestic Violence and Women's Autonomy in Developing Countries: Theory and Evidence," *Canadian Journal of Economics/Revue Canadienne d'économique*, 2011, *44* (4), 1222–63.

Farmer, Amy and Jill Tiefenthaler, *An Economic Analysis of Domestic Violence'. Review of Social Economy*, Taylor & Francis, Ltd, 1997.

Fearon, James and Anke Hoeffler, "Benefits and Costs of the Conflict and Violence Targets for the Post-2015 Development Agenda," *Conflict and violence assessment paper, Copenhagen Consensus Center*, 2014, pp. 1–65.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms," *American Economic Review*, 2021, *111* (7), 2342–75.

García-Moreno, Claudia, Cathy Zimmerman, Alison Morris-Gehring, Lori Heise, Avni Amin, Naeemah Abrahams, Oswaldo Montoya, Padma Bhate-Deosthali, Nduku Kilonzo, and Charlotte Watts, "Addressing Violence against Women: A Call to Action," *The Lancet*, 2015, *385* (9978), 1685–95.

Giuliano, Paola and Nathan Nunn, "Understanding Cultural Persistence and Change," *The Review of Economic Studies*, 2021, *88* (4), 1541–1581.

Guarnieri, Eleonora and Ana Tur-Prats, "Cultural Distance and Donflict-Related Sexual Violence," *The Quarterly Journal of Economics*, 2023, *138* (3), 1817–1861.

Gupta, Jhumka, Kathryn L. Falb, Heidi Lehmann, Denise Kpebo, Ziming Xuan, Mazeda Hossain, Cathy Zimmerman, Charlotte Watts, and Jeannie Annan, "Gender Norms and Economic Empowerment Intervention to Reduce Intimate Partner Violence against Women in Rural Côte d'Ivoire: A Randomized Controlled Pilot Study," *BMC International Health and Human Rights*, 2013, *13* (1).

Hansén, Jessica, "Political Gender Quota in Rwanda," 2017.

Haushofer, Johannes, Charlotte Ringdal, Jeremy P. Shapiro, and Xiao Yu Wang, *Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya.*, NBER Working Paper Series w25627, 2019.

Heise, L. and M. Hossain, *Measuring Intimate Partner Violence*, London: https://researchonline.lshtm.ac.uk/id/eprint/4476182/1/Measuring2017.

Hidrobo, Melissa and Lia Fernald, "Cash Transfers and Domestic Violence," *Journal of health economics*, 2013, *32* (1), 304–319.

Hossain, Mazeda, Cathy Zimmerman, T.Abramsky Ligia Kiss, Drissa Kone, Monika

Bakayoko-Topolska, Jeannie Annan, Heidi Lehmann, and Charlotte Watts, "Working with Men to Prevent Intimate Partner Violence in a Conflict-Affected Setting: A Pilot Cluster Randomized Controlled Trial in Rural Côte d'Ivoire," *BMC Public Health*, 2014, *14* (1).

Imbens, G.W. and D.B. Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences*, Cambridge University Press, 2015.

J-PAL, "The Impacts of Economic Interventions on Intimate Partner Violence," 2022. J-PAL Policy Insights. Last modified July 2022.

Kaiser-Hughes, A, M Ndangiza, and M Ikirezi, "The Rights of Women in de facto Unions to Land and Property," 2016.

Karing, Anne, "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone," *The Quarterly Journal of Economics*, 2024, *139* (4), 2083–2133.

Kerr-Wilson, Alice, Andrew Gibbs, E McAslan Fraser, Leanne Ramsoomar, Anna Parke, Hussaind MA Khuwaja, and Rachel Jewkes, "A Rigorous Global Evidence Review of Interventions to Prevent Violence against Women and Girls," *What Works to Prevent Violence among Women and Girls Global Programme, Pretoria, South Africa*, 2020.

Klugman, Jeni, Lucia Hanmer, Sarah Twigg, Tazeen Hasan, Jennifer McCleary-Sills, and Julieth Santamaria, *Voice and Agency: Empowering Women and Girls for Shared Prosperity*, The World Bank, 2014.

LaMattina, Giulia, "Civil Conflict, Domestic Violence and Intra-Household Bargaining in Post-Genocide Rwanda," *Journal of Development Economics*, 2017, *124*, 168–98.

Mannell, Jenevieve, Iran Seyed-Raeisy, Rochelle Burgess, and Catherine Campbell, "The Implications of Community Responses to Intimate Partner Violence in Rwanda," *PLOS ONE*, 2018, *13* (5), 0196584.

McIntosh, Craig and Andrew Zeitlin, "Cash versus kind: Benchmarking a child nutrition program against unconditional cash transfers in Rwanda," *The Economic Journal*, 2024, p. ueae050.

Mulla, Mazheruddin M., Tricia H. Witte, Kyle Richardson, William Hart, Francesca L. Kassing, Christopher A. Coffey, Christine L. Hackman, and Ian M. Sherwood, "The Causal Influence of Perceived Social Norms on Intimate Partner Violence Perpetration: Converging Cross-Sectional, Longitudinal, and Experimental Support for a Social Disinhibition Model," *Personality and Social Psychology Bulletin*, 2019, *45* (4), 652–668. PMID: 30226406.

Oh, Suanna, "Does Identity Affect Labor Supply?," *American Economic Review*, 2023, *113* (8), 2055–2083.

Roy, Shalini, Melissa Hidrobo, John Hoddinott, and Akhter Ahmed, "Transfers, Behavior Change Communication, and Intimate Partner Violence: Postprogram Evidence from Rural Bangladesh," *Review of Economics and Statistics*, 2019, *101* (5), 865–877.

Sanin, Deniz, "Do Domestic Violence Laws Protect Women From Domestic Violence? Evidence From Rwanda," 2021.

_ , "Women's Employment, Husbands' Economic Self-Interest and Domestic Violence," 2023.

Shah, Manisha, Jennifer Seager, Joao Montalvao, and Markus Goldstein, "Sex, Power, and Adolescence: Intimate Partner Violence and Sexual Behaviors," Technical Report, National Bureau of Economic Research 2023.

Sharma, Vandana, Jessica Leight, Fabio Verani, Samuel Tewolde, and Negussie Deyessa, "Effectiveness of a Culturally Appropriate Intervention to Prevent Intimate Partner Violence and HIV Transmission among Men, Women, and Couples in Rural Ethiopia: Findings from a Cluster-Randomized Controlled Trial," *PLoS Medicine*, 2020, *17* (8), 1–28.

Slegh, Henny, Gary Barker, Augustin Kimonyo, Prudence Ndolimana, and Matt Bannerman, ""I Can Do Women's Work": Reflections on Engaging Men as Allies in Women's Economic Empowerment in Rwanda," *Gender & Development*, 2013, *21* (1), 15–30.

Stern, Erin and Ritha Nyiratunga, "A Process Review of the Indashyikirwa Couples Curriculum to Prevent Intimate Partner Violence and Support Healthy, Equitable Relationships in Rwanda," *Social Sciences*, 2017, *6* (2), 63.

Tankard, Margaret E. and Elizabeth Levy Paluck, "Norm Perception as a Vehicle for Social Change," *Social Issues and Policy Review*, 2016, *10* (1), 181–211.

Tauchen, Helen V., Ann Dryden Witte, and Sharon K. Long, "Domestic Violence: A Nonrandom Affair," *International Economic Review*, 1991, *32* (2), 491.

Tur-Prats, Ana, "Family Types and Intimate Partner Violence: A Historical Perspective," *Review of Economics and Statistics*, 2019, *101* (5), 878–891.

USAID, "Final Report - Rwanda Land Project," 2016.

Vaillant, J., E. Koussoubé, D. Roth, R. Pierotti, M. Hossain, and K. Falb, "Engaging Men to Transform Inequitable Gender Attitudes and Prevent Intimate Partner Violence: A Cluster Randomised Controlled Trial in North and South Kivu," *BMJ Global Health*, 2020, *5* (5).

WHO, "Ethical and Safety Recommendations for Intervention Research on Violence against Women'," 2016.

Tables

Table 1: Treatment effects on primary IPV outcomes

	(1) Emotional IPV	(2) Physical IPV	(3) Sexual IPV
Treated couple	0.07**	0.05*	0.10***
	(0.03)	(0.02)	(0.03)
	[0.007]	[0.014]	[0.001]
Spillover couple	0.06***	0.11***	0.17***
	(0.02)	(0.02)	(0.02)
	[0.004]	[0.001]	[0.001]
Observations	1,974	1,960	1,972
Baseline controls	Yes	Yes	Yes
Control mean	0.66	0.33	0.42
Treated vs Spillover p-value	0.92	0.02	0.01

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in (rounded brackets). Sharpened q-values are shown in [square brackets].
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples.
Sharpened q-values were calculated to adjust for multiple hypothesis testing using the Benjamini and Hochberg (1995) correction method.
The unit of observation is individual level, and is calculated using a dummy aggregating whether a wife reported experiencing any of each of the forms of emotional, physical, or sexual violence, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Control mean is the average value of the dependent variable in the control group.

Table 2: Treatment effects on other well-being and relationship quality indicators

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Relationship quality		Depression index		Disagreement escalated into violence	Husband drunk before last quarrel	Husband takes money for himself	Husband cheated on wife
	Women	Men	Women	Men				
Treated couple	-0.05	-0.01	-0.02	0.03	0.09***	0.05*	0.06**	0.02
	(0.04)	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)
Spillover couple	-0.19***	-0.18***	0.10**	0.15***	0.07***	0.07**	0.09***	0.07***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)
Observations	2,042	2,042	2,042	2,042	1,978	1,965	2,042	1,734
Baseline controls	No	No	No	No	No	No	No	No
Control mean	0.00	0.00	0.00	0.00	0.10	0.36	0.37	0.12
Treated vs Spillover p-value	0.00	0.00	0.00	0.01	0.30	0.62	0.44	0.03

Note: OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, not controlling for baseline values of dependent variables.
Columns 1–4 use a weighted average of relationship quality and mental health variables. Columns 5–7 use dichotomous variables based on wives' answers.
In column 5, the dependent variable is equal to one if the last marital disagreement reportedly escalated into violence. In column 6, it is equal to one if the husband was drunk before the last quarrel. In column 7, it is equal to one if the wife reports her husband spends, on himself, money set aside for the household. In column 8, it is equal to one if the wife reports that her husband cheated on her.
Control mean is the average value of the dependent variable in the control group.

Table 3: Backlash mechanisms: Behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Wife reports husband drinks too much	Drank alcohol to excess in past month Men	Number of drinks in past month Men	Formalized marriage	Husband took household money for himself	Minutes doing domestic or care work as primary activity	
						Women	Men
Treated couple	0.07***	0.04	0.25	0.24***	0.06**	17.83**	12.34***
	(0.02)	(0.04)	(0.41)	(0.03)	(0.03)	(8.30)	(4.40)
Spillover couple	0.12***	0.08**	0.86**	-0.00	0.09***	13.63	5.70
	(0.02)	(0.04)	(0.36)	(0.02)	(0.03)	(8.46)	(4.41)
Observations	2,030	2,032	2,042	1,003	2,042	2,042	2,042
Baseline controls	Yes	Yes	No	No	No	No	No
Control mean	0.39	0.50	4.06	0.04	0.37	270.70	41.49
Treated vs Spillover p-value	0.05	0.35	0.13	0.00	0.44	0.66	0.20

Note: OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples.
Columns 1 and 2 control for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Columns 2 and 3 use variables based on men's self-reported alcohol consumption. Column 4 restricts the sample to the couples who were not legally married at baseline and uses a dichotomous variable equal to one if they formalized their marriage between baseline and endline. Column 5 uses a dichotomous variable equal to one if the wife reports her husband spent, on himself, money set aside for the household.
Control mean is the average value of the dependent variable in the control group.

Table 4: Backlash mechanisms: Agency and decision-making

Panel A: Women	(1)	(2)	(3)	(4)	(5)	(6)
	Want more decision-making on		Has input in decisions regarding			
	Household expenses	Childbearing	Income	Household income	Major household expenses	Childbearing
Treated couple	0.06***	0.05**	0.07***	0.06**	0.06*	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
Spillover couple	0.09***	0.01	-0.03	-0.05*	-0.02	0.00
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
Observations	1,984	1,775	1,975	1,986	1,987	1,760
Baseline controls	No	No	No	No	No	No
Control mean	0.21	0.14	0.72	0.58	0.60	0.84
Treated vs Spillover p-value	0.21	0.10	0.00	0.00	0.06	0.50
Panel B: Men						
	Want more decision-making on		Is sole decision-maker on			
	Household expenses	Childbearing	Income	Household income	Major household expenses	Childbearing
Treated couple	-0.00	0.01	-0.01	-0.01	0.00	-0.01*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Spillover couple	0.03*	0.02	0.04**	0.02**	0.03**	-0.00
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Observations	1,978	1,728	1,969	1,979	1,978	1,746
Baseline controls	No	No	No	No	No	No
Control mean	0.07	0.05	0.07	0.02	0.03	0.03
Treated vs Spillover p-value	0.05	0.53	0.00	0.00	0.07	0.35

Note: OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, not controlling for baseline values of dependent variables.
Columns 1–2 use dichotomous variables equal to one if the spouse declares wanting more decision-making power. Columns 3–6 use dichotomous variables equal to one if the wife declares having a role in the decision (panel A) or the husband declares being the sole decision-maker (panel B).
Control mean is the average value of the dependent variable in the control group.

Table 5: Backlash mechanisms: Gender attitudes

Panel A: Women	(1) No. gender inequitable attitudes	(2) No. situations thinks IPV is justified	(3) No. situations refusing sex is justified
Treated couple	-0.40***	-0.60***	-0.19***
	(0.04)	(0.08)	(0.06)
Spillover couple	-0.10**	-0.23***	-0.14**
	(0.04)	(0.08)	(0.06)
Observations	2,008	1,732	1,815
Baseline controls	Yes	Yes	Yes
Control mean	1.82	2.19	1.90
Treated vs Spillover p-value	0.00	0.00	0.46

Panel B: Men	(1) No. gender inequitable attitudes	(2) No. situations thinks IPV is justified	(3) No. situations wife refusing sex is justified	(4) No. hypothetical situations would be likely to become violent
Treated couple	-0.27***	-0.22***	-0.10*	-0.41**
	(0.05)	(0.06)	(0.06)	(0.19)
Spillover couple	0.10	0.30***	-0.08	0.95***
	(0.06)	(0.08)	(0.06)	(0.23)
Observations	1,989	1,640	1,746	1,990
Baseline controls	Yes	Yes	Yes	Yes
Control mean	1.61	0.83	2.03	3.46
Treated vs Spillover p-value	0.00	0.00	0.81	0.00

Panel C: Frictions	(1)	(2)	(3)
	Difference husband - wife		
	Gender inequitable attitudes	IPV acceptability attitudes	Wife got more progressive Husband got more conservative
Treated couple	0.19***	0.43***	-0.01
	(0.06)	(0.09)	(0.01)
Spillover couple	0.15**	0.61***	0.05***
	(0.06)	(0.09)	(0.02)
Observations	2,042	2,042	2,042
Baseline controls	Yes	Yes	No
Control mean	-0.49	-1.19	0.09
Treated vs Spillover p-value	0.57	0.13	0.00

Note: OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
panel A reports the treatment effect on women's answers; panel B reports on men's answers. Column 1 uses the number of gender inequitable attitudes declared by the respondent (out of 4). Column 2 uses the number of reasons that justify IPV (out of 5). Column 3 uses the number of situations in which refusing sex is justified (out of 4). Column 4 uses the number of hypothetical situations in which a man would beat his wife (out of 10) or be intimidating (out of 5).
Columns 1 and 2 of panel C use the difference between in gender inequitable attitudes and the number of situations IPV is justified between spouses. A positive value implies that the wife is more progressive than the husband. Column 3 uses a dichotomous variable equal to one if the wife became more progressive between baseline and endline while her husband became more conservative.
Control mean is the average value of the dependent variable in the control group.

Figures

Figure 2: Experimental design

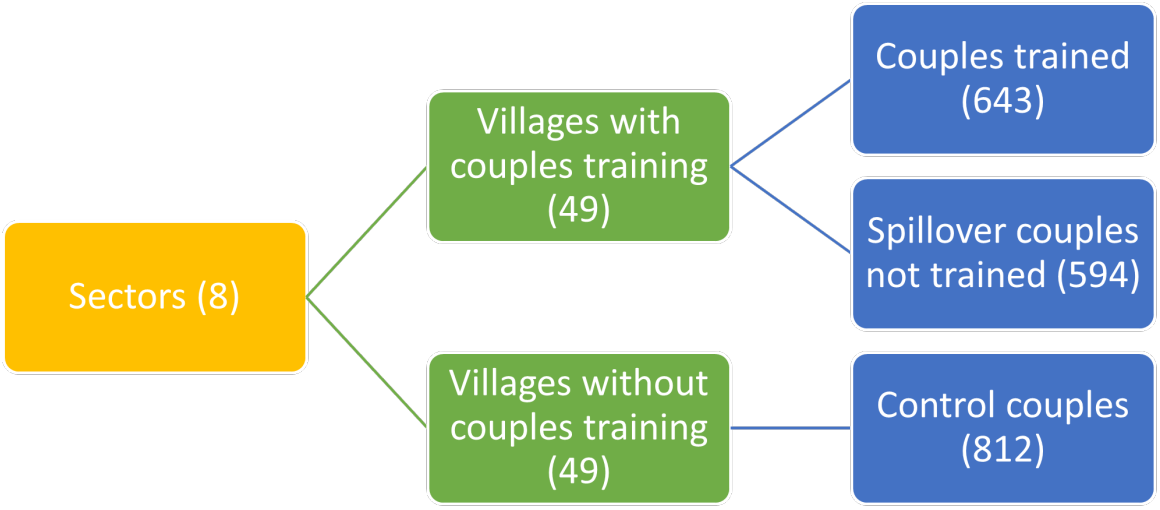
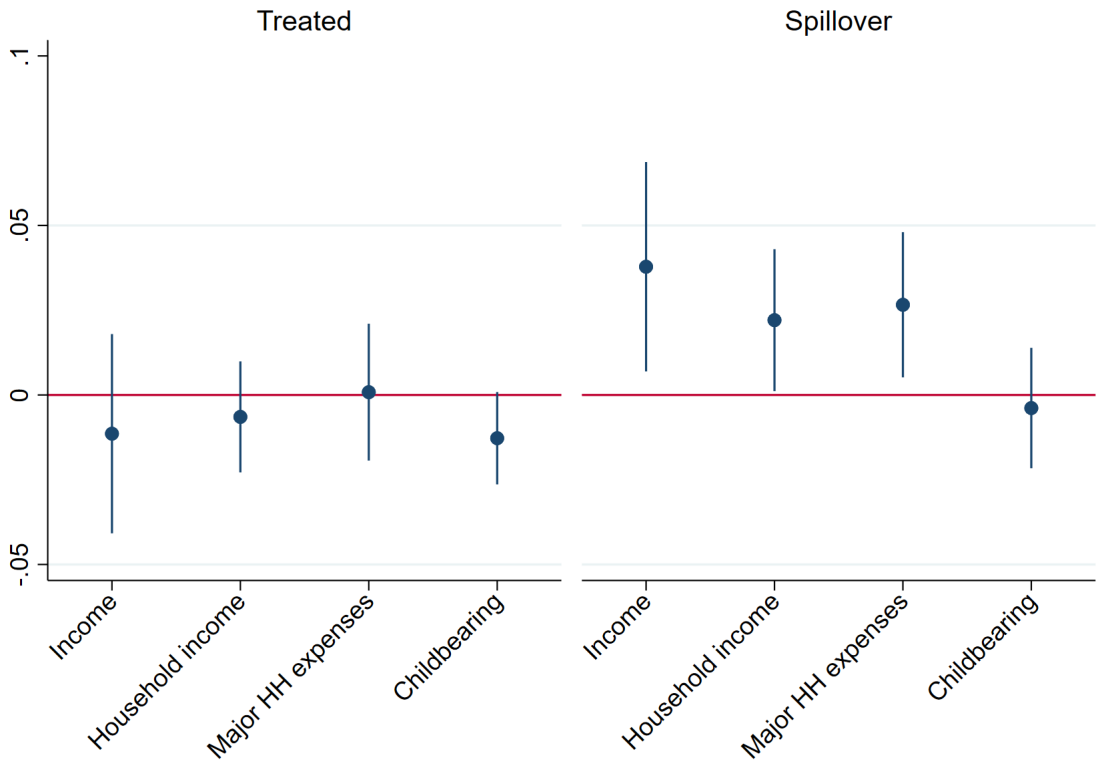
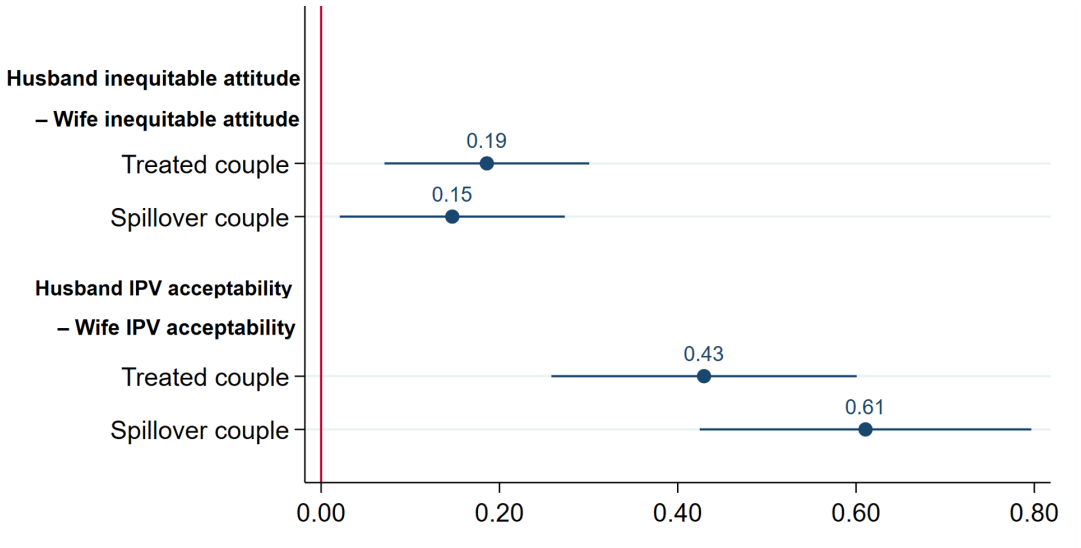


Figure 3: Treatment effects on agency: Husband is the sole decision-maker



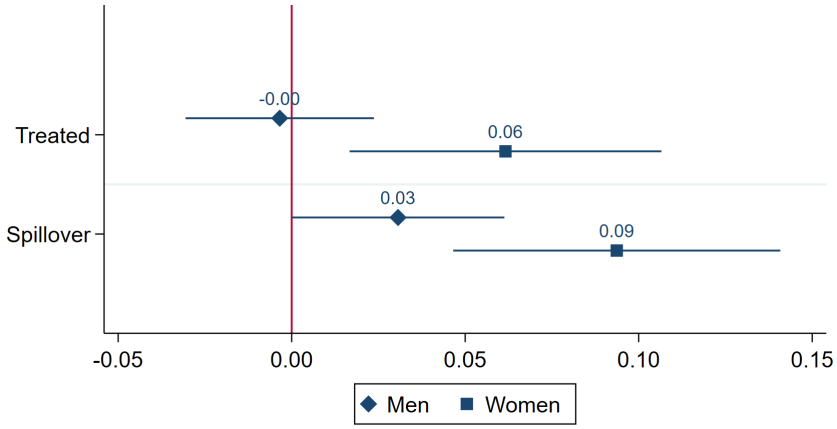
Note: HH = household. Treatment and spillover effects on attitude divergence betweeen spouses, with 95 percent confidence intervals (table 5, panel C). Separate regression for each outcome variable.

Figure 4: Treatment effects on attitude divergence between spouses



Note: IPV = intimate partner violence. Treatment and spillover effects on attitude divergence between spouses, with 95 percent confidence intervals (table 5, panel C). Outcome variable written in bold, separate regression for each outcome variable.

Figure 5: Treatment effects on agency aspirations: Wants more decision-making on household expenses



Note: Treatment and spillover effects on husband's sole decision making, with 95 percent confidence intervals (table 4 panel B). Separate regression for each decision domain.

Appendix A Additional Tables and Figures

Table A1: Couples' curriculum

Session #1: Starting the Journey Together
- Part 1: Who We Are and Why We're Here
- Part 2: Great Expectations
- Part 3: Creating a Safe Space
- Part 4: Introducing the Take-Home Exercise

Session #2: How We Change
- Part 1: Stages of Change
- Part 2: Motivators and Barriers to Change

Session #3: It's All a Power
- Part 1: Types of Power
- Part 2: Power and Indashyikirwa

Session #4: Power in Our Lives
- Part 1: Our Experiences of Power
- Part 2: Who Has Power?

Session #5: G is for Gender
- Part 1: Gender Lifelines
- Part 2: The 24-Hour Day

Session #6: Rights and Reality
- Part 1: Human Rights and Reality
- Part 2: What the Law Says

Session #7: GBV- The Basics
- Part 1: What Is Gender-Based Violence?
- Part 2: Effects of GBV
- Part 2: Attitudes

Session #8: Understanding 'Power Over'
- Part 1: The New Planet
- Part 2: 'Power Over' in Our Relationships

Session #9: Gender, Power, and Sexuality
- Part 1: Double Standards
- Part 2: The Truth about Our Sexuality

Session #10: Common Triggers of GBV in Couples
- Part 1: Identifying Triggers
- Part 2: Understanding Triggers

Session #11: The Children Are Our Future
- Part 1: The Impact of GBV on Children
- Part 2: A Better Future for Our Children

Session #12: Pausing for Reflection
- Part 1: Participant Teach-Back
- Part 2: Personal Reflection
- Part 3: Skills

Session #13: What Makes a Healthy Relationship?
- Part 1: Characteristics of a Healthy Relationship
- Part 2: Feelings into Actions

Session #14: Building the Foundations of a Healthy Relationship
- Part 1: Sharing Positive Time
- Part 2: The Power of Communication

Session #15: Managing Triggers of GBV - Akarimi Koshywa N'Akandi
- Part 1: What We Say and How We Say It
- Part 2: Conversation in a Conflict

Session #16: Balancing Economic Power
- Part 1: Economic Power in Our Relationships
- Part 2: The Gender and VSL 'Fishbowl'

Session #17: Reducing Excessive Use of Alcohol
- Part 1: Throwing the Drunk Ball
- Part 2: Strategies for Reducing Alcohol Consumption

Session #18: Reflecting on Our Journey So Far
- Part 1: Participant Teach-Back
- Part 2: Personal Reflection
- Part 4: Action

Session #19: Our Community, Our Responsibility
- Part 1: Circles of Influence
- Part 2: Support or Interference

Session #20: Providing Empowering Response
- Part 1: Moving beyond Shame and Stigma
- Part 2: Fostering 'Power Within'

Session #21: Small Actions, Big Impact
- Part 1: Instruct, Inform, or Question
- Part 2: Starting the Conversation

Session #22: Committing to Change!
- Part 1: Roles of an Activist
- Part 2: Making Commitments

Table A2: Baseline differences between VSLA couple applicant sample and random community couple sample

Variable	(1) Community Couple N/Clusters	(1) Community Couple Mean/(SE)	(2) VSLA Couple N/Clusters	(2) VSLA Couple Mean/(SE)	(1) − (2) Pairwise t-test t-statistics
Wife variables					
Any emotional violence in last 12m	674	0.49	2006	0.48	0.48
	98	(0.02)	98	(0.01)	
Any physical violence in last 12m	666	0.40	1987	0.39	0.03
	98	(0.02)	98	(0.01)	
Any sexual violence in last 12m	658	0.45	1988	0.43	0.71
	98	(0.02)	98	(0.01)	
Age	686	38.98	2042	36.60	4.09***
	98	(0.38)	98	(0.34)	
Years of education	686	3.89	2042	4.51	-4.04***
	98	(0.14)	98	(0.09)	
No. of children	686	3.94	2041	3.73	2.22**
	98	(0.08)	98	(0.06)	
No. of reasons IPV is justified	546	2.52	1676	2.30	2.22**
	98	(0.09)	98	(0.06)	
No. of gender inequitable attitudes	684	2.00	2039	2.01	0.22
	98	(0.03)	98	(0.02)	
Husband variables					
Age	686	43.22	2042	40.27	6.50***
	98	(0.41)	98	(0.29)	
Years of education	686	4.10	2042	4.69	-3.79***
	98	(0.14)	98	(0.08)	
No. of reasons IPV is justified	553	1.14	1663	0.96	2.29**
	98	(0.07)	98	(0.03)	
No. of gender inequitable attitudes	683	1.54	2039	1.50	0.91
	98	(0.03)	98	(0.02)	
Thinks at most 25% of men beat up wife	686	0.90	2042	0.92	-1.57
	98	(0.01)	98	(0.01)	
Thinks at most 25% of men force wife to have sex	686	0.88	2042	0.90	-1.73*
	98	(0.01)	98	(0.01)	
Couple variables					
In a civil or legal marriage	686	0.49	2042	0.50	-0.47
	98	(0.02)	98	(0.02)	
A member of HH skipped a meal in the last week	686	0.45	2042	0.40	1.99**
	98	(0.02)	98	(0.02)	

Note: HH = household; IPV = intimate partner violence; m = months; N = number; No. = number; SE = standard error; VSLA = village savings and loans association. The normalized differences reflect differences in the means across groups.
Standard errors are clustered at the village level. Fixed effects using sector and number of program applicant strata are included in all estimations.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A3: Baseline Sample Characteristics and Balance across three groups

	(1)		(2)		(3)		(1) − (2)	(1) − (3)	(2) − (3)
	Pure control		Spillover couple		Treated couple		Pairwise t-test		
Variable	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)	N/Clusters	Mean/(SE)			
Wife variables									
Any emotional violence in last 12m	798	0.47	583	0.48	625	0.51	-0.52	-1.48	-0.96
	49	(0.02)	49	(0.02)	49	(0.02)			
Any physical violence in last 12m	793	0.40	575	0.38	619	0.40	0.99	0.09	-0.83
	49	(0.02)	49	(0.02)	49	(0.02)			
Any sexual violence in last 12m	795	0.45	575	0.40	618	0.44	1.48	0.48	-1.32
	49	(0.02)	49	(0.02)	49	(0.02)			
Age	812	36.07	594	36.84	636	37.07	-1.66*	-1.43	-0.38
	49	(0.41)	49	(0.41)	49	(0.80)			
Years of education	812	4.57	594	4.52	636	4.44	0.54	1.10	0.48
	49	(0.14)	49	(0.13)	49	(0.16)			
No. of children	812	3.68	594	3.80	635	3.72	-1.15	-0.72	0.72
	49	(0.09)	49	(0.10)	49	(0.09)			
No. of reasons IPV is justified	663	2.40	489	2.24	524	2.22	1.80*	1.63	-0.02
	49	(0.08)	49	(0.10)	49	(0.10)			
No. of gender inequitable attitudes	811	2.03	593	1.93	635	2.05	2.35**	-0.29	-2.06**
	49	(0.03)	49	(0.04)	49	(0.04)			
Husband variables									
Age	812	40.05	594	40.45	636	40.39	-0.79	-0.78	0.06
	49	(0.45)	49	(0.46)	49	(0.47)			
Years of education	812	4.79	594	4.57	636	4.67	1.87*	1.12	-0.61
	49	(0.12)	49	(0.13)	49	(0.13)			
No. of reasons IPV is justified	652	0.93	490	0.88	521	1.06	0.43	-2.04**	-2.59**
	49	(0.05)	49	(0.06)	49	(0.06)			
No. of gender inequitable attitudes	811	1.47	594	1.57	634	1.47	-1.97*	0.08	1.81*
	49	(0.03)	49	(0.05)	49	(0.04)			
Thinks at most 25% of men beat up wife	812	0.92	594	0.92	636	0.93	0.06	-0.37	-0.27
	49	(0.01)	49	(0.01)	49	(0.01)			
Thinks at most 25% of men force wife to have sex	812	0.90	594	0.89	636	0.91	0.59	-0.62	-0.92
	49	(0.01)	49	(0.01)	49	(0.01)			
Couple variables									
In a civil or legal marriage	812	0.51	594	0.47	636	0.51	1.37	-0.02	-1.58
	49	(0.03)	49	(0.03)	49	(0.03)			
Member of household skipped meal in last week	812	0.43	594	0.39	636	0.39	1.01	1.33	0.05
	49	(0.03)	49	(0.03)	49	(0.03)			

Note: IPV = intimate partner violence; m = months; N = number; No. = number; SE = standard error. The normalized differences reflect differences in the means across groups. Standard errors are clustered at the village level. Fixed effects using sector and number of program applicant strata are included in all estimation regressions.

Table A4: Treatment effects excluding respondents who live in unconfirmed or different village to VSLA-village of treatment assignment

	(1) Emotional IPV	(2) Physical IPV	(3) Sexual IPV
Treated couple	0.10***	0.07**	0.08**
	(0.03)	(0.03)	(0.04)
Spillover couple	0.09***	0.13***	0.15***
	(0.03)	(0.03)	(0.04)
Observations	1,097	1,088	1,098
Baseline controls	Yes	Yes	Yes

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Sample restricted to households for whom it was possible to confirm that they live in the VSLA-village of treatment assignment.

Table A5: Treatment effects excluding control couples living near treated and spillover couples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Emotional IPV				Physical IPV				Sexual IPV			
	200 m exclusion	400 m exclusion	600 m exclusion	800 m exclusion	200 m exclusion	400 m exclusion	600 m exclusion	800 m exclusion	200 m exclusion	400 m exclusion	600 m exclusion	800 m exclusion
Treated couple	0.05**	0.05*	0.06*	0.06*	0.06**	0.06**	0.06**	0.07***	0.10***	0.10***	0.11***	0.10***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Spillover couple	0.05**	0.05**	0.05**	0.05*	0.12***	0.12***	0.13***	0.14***	0.17***	0.17***	0.17***	0.16***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Observations	1,889	1,819	1,742	1,662	1,873	1,805	1,730	1,649	1,887	1,817	1,741	1,660
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: IPV = intimate partner violence; m = meters; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Columns 1, 5, and 9 exclude control couples living less than 200 m away from treated or spillover couples, as well as treated or spillover couples living more than 200 m away from any other treated or spillover couples. Columns 2, 6, and 10 use a distance threshold of 400 m; columns 3, 7, and 11 use a threshold of 600 m; and columns 4, 8, and 12 use a threshold of 800 m.

Table A6: Spillover treatment-on-the-treated estimates — instrumenting spillovers' village residence

	(1)	(2)	(3)	(4)	(5)	(6)
	Emotional IPV		Physical IPV		Sexual IPV	
	First stage	Second stage	First stage	Second stage	First stage	Second stage
Spillover couple assignment	0.65***		0.64***		0.65***	
	(0.03)		(0.03)		(0.03)	
Spillover couple complier		0.12***		0.19***		0.27***
		(0.04)		(0.04)		(0.04)
Observations	1,126	1,126	1,111	1,111	1,123	1,123
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes

Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: IPV = intimate partner violence; VSLA = village savings and loans association; 2SLS = two-stage least square. Results show 2-SLS estimates of direct treatment effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing. VSLA-village treatment assignment is used as instrument for actual treatment.

Table A7: Treatment effects on IPV-related outcomes

Panel A: Women	(1) Experience of husband's controlling behaviours in last 6m	(2) Reported having bruises sprains, etc. from IPV	(3) Reported having injuries from sexual IPV	(4) Enumerator: Woman shows signs of physical pain or injuries	(5) Used physical violence against children in past 6m	(6) Last marital disagreement about sex
Treated couple	0.02	0.03*	0.06***	0.00	-0.03	0.02**
	(0.03)	(0.02)	(0.02)	(0.00)	(0.03)	(0.01)
Spillover couple	0.11***	0.06***	0.08***	0.01**	0.00	0.02**
	(0.03)	(0.02)	(0.02)	(0.00)	(0.03)	(0.01)
Observations	1,919	2,042	2,042	2,008	1,968	1,987
Baseline controls	Yes	Yes	Yes	Yes	No	No
Control mean	0.59	0.11	0.08	0.00	0.67	0.02
Treated vs Spillover p-value	0.00	0.20	0.27	0.28	0.22	0.56

Panel B: Men	(1) Experienced emotional IPV in past 6m	(2) Experienced physical IPV in past 6m	(3) Frequency of physical violence in village in past 6 months	(4) Frequency of physical violence against women in village in past 6 months
Treated couple	0.08***	0.00	0.46**	0.39**
	(0.03)	(0.02)	(0.18)	(0.18)
Spillover couple	0.04	0.07***	0.37*	0.38*
	(0.03)	(0.02)	(0.21)	(0.22)
Observations	1,965	1,959	1,991	1,486
Baseline controls	Yes	Yes	No	No
Control mean	0.60	0.29	2.42	2.56
Treated vs Spillover p-value	0.17	0.01	0.70	0.97

Note: IPV = intimate partner violence; m = months; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples.
Panel A shows the results for IPV-related outcomes reported by women, except for column 4, where it is a variable observed by the enumerator.
Panel B shows the results for outcomes reported by men. Columns 3 and 4 use the total number of violent events in the village, as reported by the men, with the top 1 percent winsorized.
Columns 1–4 of panel A and 1–2 of panel B control for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Control mean is the average value of the dependent variable in the control group.

Table A8: Treatment effect on economic outcomes

	(1) Number of assets HH owns (out of 6)	(2) Share of HH assets owned by wife	(3) Involved in revenue-generating activity in past 12m	(4)	(5) Savings	(6)
			Women	Men	Women	Men
Treated couple	0.31*	0.00	-0.01	-0.01	8,127.89***	4,829.77
	(0.16)	(0.01)	(0.03)	(0.02)	(2,675.69)	(4,879.37)
Spillover couple	0.30*	-0.02	-0.03	-0.01	2,470.59	1,180.05
	(0.17)	(0.01)	(0.03)	(0.02)	(2,327.91)	(5,562.96)
Observations	2,042	2,007	2,008	1,991	2,008	1,991
Baseline controls	No	No	No	No	Yes	Yes
Control mean	5.82	0.30	0.46	0.38	31,201	57,924.2
Treated vs Spillover p-value	0.97	0.24	0.46	0.82	0.01	0.56

Note: HH = household; m = months; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples.
Columns 5–6 control for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing. The other columns do not.
Columns 1 and 2 show results for asset ownership, out of six types of assets (radio, phone, scooter, bicycle, television, and mats).
Columns 3 and 4 show results for a dichotomous variable equal to one if the respondent was involved in a revenue-generating activity during the past year.
Columns 5 and 6 use the amount of savings on the respondent's main savings instrument account, with the top 1 percent winsorized.
Control mean is the average value of the dependent variable in the control group.

Table A9: Heterogeneous treatment effects by baseline violence experience

	(1) Emotional IPV	(2) Physical IPV	(3) Sexual IPV
Treated couple	0.06**	0.03	0.09***
	(0.03)	(0.03)	(0.03)
Spillover couple	0.01	0.07**	0.13***
	(0.03)	(0.03)	(0.03)
Treated * No violence at baseline	0.03	0.03	0.01
	(0.06)	(0.04)	(0.05)
Spillover * No violence at baseline	0.17***	0.14**	0.11*
	(0.06)	(0.06)	(0.06)
No violence at baseline	-0.24***	-0.12***	-0.14***
	(0.04)	(0.03)	(0.04)
Observations	1,936	1,923	1,937
Baseline controls	Yes	Yes	Yes
P-value (Treated + Treated * no violence = 0)	0.09	0.08	0.02
P-value (Spillover + Spillover * no violence = 0)	0.00	0.00	0.00

Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Note: IPV = intimate partner violence; OLS = ordinary least squares. Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
No violence at baseline is a dichotomous variable equal to one if the woman does not report experiencing any type of IPV at baseline.
The last two rows show the p-value of the treatment and spillover effects on women who did not experience any violence at baseline.

Table A10: Impact by IPV question (subjective and objective)

Panel A: Emotional IPV

	(1) Publicly humiliated	(2) Insulted
Treated couple	0.10***	0.05**
	(0.03)	(0.03)
Spillover couple	0.11***	0.06***
	(0.02)	(0.02)
Observations	1,962	1,966
Baseline controls	Yes	Yes
Control mean	0.30	0.60
Treated vs Spillover p-value	0.91	0.76
More subjective interpretation in Kinyarwanda	Yes	No

Panel B: Physical IPV

	(1) Slapped	(2) Twisted arm	(3) Punched	(4) Kicked or beat	(5) Strangled or burned	(6) Threatened to attack with weapon or knife
Treated couple	0.04*	0.04*	0.06***	0.04**	-0.01	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Spillover couple	0.09***	0.05***	0.09***	0.09***	0.00	0.03**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Observations	1,977	1,968	1,974	1,981	1,969	1,972
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.24	0.15	0.14	0.15	0.09	0.06
Treated vs Spillover p-value	0.05	0.39	0.20	0.04	0.49	0.42
More subjective interpretation in Kinyarwanda	No	No	No	No	No	No

Panel C: Sexual IPV

	(1) Forced to have sex when she did not want	(2) Forced to have sex when she did was afraid of him	(3) Forced to do degrading sex acts
Treated couple	0.09***	0.05*	0.03*
	(0.02)	(0.02)	(0.02)
Spillover couple	0.15***	0.14***	0.04*
	(0.02)	(0.03)	(0.02)
Observations	1,970	1,982	1,962
Baseline controls	Yes	Yes	Yes
Control mean	0.35	0.27	0.15
Treated vs Spillover p-value	0.03	0.00	0.87
More subjective interpretation in Kinyarwanda	No	No	Yes

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Each column presents the results for a specific type of IPV, reported by the woman.
Control mean is the average value of the dependent variable in the control group.

Table A11: Treatment effects on increase in frequency and number of types of IPV

Panel A: Physical IPV : Sometimes or often	(1) Slapped	(2) Twisted arm	(3) Punched	(4) Kicked or beat	(5) Strangled or burned	(6) Threatened to attack with weapon or knife
Treated couple	0.04**	0.01	0.03*	0.03**	0.00	0.02*
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Spillover couple	0.09***	0.03*	0.04**	0.07***	0.03**	0.03***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	1,977	1,968	1,974	1,981	1,969	1,972
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.11	0.08	0.08	0.06	0.03	0.01
Treated vs Spillover p-value	0.05	0.15	0.38	0.01	0.05	0.21

Panel B: Sexual IPV : Sometimes or often	(1) Forced to have sex when she did not want	(2) Forced to have sex when she was afraid of him	(3) Forced to do degrading sex acts
Treated couple	0.05**	0.03	-0.00
	(0.02)	(0.02)	(0.01)
Spillover couple	0.09***	0.10***	0.02
	(0.02)	(0.02)	(0.02)
Observations	1,970	1,982	1,962
Baseline controls	Yes	Yes	Yes
Control mean	0.21	0.15	0.08
Treated vs Spillover p-value	0.20	0.01	0.15

Panel C: Number of types of violence	(1) Emotional IPV (out of 2)	(2) Physical IPV (out of 6)	(3) Sexual IPV (out of 3)
Treated couple	0.15***	0.18**	0.17***
	(0.05)	(0.08)	(0.06)
Spillover couple	0.16***	0.36***	0.32***
	(0.04)	(0.08)	(0.06)
Observations	2,008	2,008	2,008
Baseline controls	Yes	Yes	Yes
Control mean	0.88	0.81	0.77
Treated vs Spillover p-value	0.92	0.03	0.01

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Each columns of panels A and B presents the results for a specific type of IPV, using a dichotomous dependent variable equal to one if the woman reports experiencing the type of IPV "sometimes" or "often" (as opposed to "never" or "rarely").
Panel C presents the results using the number of different IPV types experienced by the woman.
Control mean is the average value of the dependent variable in the control group.

Table A12: Treatment effect on refusal to answer IPV-related questions

	(1)	(2)	(3)
	Refused to answer at least one question related to		
	Emotional IPV	Physical IPV	Sexual IPV
Treated couple	-0.00	-0.02*	-0.00
	(0.01)	(0.01)	(0.01)
Spillover couple	0.01	0.03**	0.01
	(0.01)	(0.01)	(0.01)
Observations	2,042	2,042	2,042
Baseline controls	Yes	Yes	Yes
Control mean	0.04	0.06	0.04
Treated vs Spillover p-value	0.27	0.00	0.59

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
The dependent variables are dichotomous variables equal to one if the woman refused to answer at least one question related to the IPV category.
Control mean is the average value of the dependent variable in the control group.

Table A13: Treatment effects on indicators of stigma and empowerment to report

	(1)	(2)	(3)	(4)
	Women			Men
	Visited safe space in past 6 months	Spoken with others about physical or sexual IPV suffered	Would report IPV witnessed to local leaders (hypothetical)	Visited safe space in past 6 months
Treated couple	0.07**	0.06*	-0.04*	0.04**
	(0.03)	(0.03)	(0.02)	(0.02)
Spillover couple	-0.02	0.02	-0.03	-0.01
	(0.03)	(0.03)	(0.02)	(0.02)
Observations	2,008	1,586	1,013	1,991
Baseline controls	No	No	No	No
Control mean	0.22	0.27	0.12	0.11
Treated vs Spillover p-value	0.00	0.34	0.76	0.02

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, not controlling for the baseline value of dependent variable.
Columns 1–3 use variables reported by women, column 4 uses a variable reported by men.
Control mean is the average value of the dependent variable in the control group.

Table A14: Correlations between women's IPV experience and potentially objective indicators

	(1)						
	Emotional IPV	Physical IPV	Sexual IPV	Wife shows signs of distress	Wife shows signs of physical pain or injuries	Wife was referred to counselor	Wife chose to speak to counselor
Emotional IPV	1.00						
Physical IPV	0.38***	1.00					
Sexual IPV	0.38***	0.40***	1.00				
Wife shows signs of distress	0.06**	0.08***	0.08***	1.00			
Wife shows signs of physical pain or injuries	-0.02	0.01	0.02	0.10***	1.00		
Wife was referred to counselor	0.05*	0.08***	0.08***	0.85***	0.30***	1.00	
Wife chose to speak to counselor	0.09***	0.11***	0.09***	0.43***	0.06*	0.48***	1.00
Observations	1,916						

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: IPV = intimate partner violence.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.5$

Table A15: Heterogeneous effects by husband migration status

	(1) Emotional IPV	(2) Physical IPV	(3) Sexual IPV
Treated couple	0.06*	0.06*	0.10**
	(0.04)	(0.03)	(0.04)
Spillover couple	0.07*	0.14***	0.19***
	(0.04)	(0.03)	(0.04)
Treated * Husband worked away	0.03	-0.04	0.04
	(0.09)	(0.08)	(0.08)
Spillover * Husband worked away	-0.04	-0.19**	-0.16**
	(0.10)	(0.08)	(0.08)
Husband worked away	-0.02	0.13**	-0.00
	(0.07)	(0.05)	(0.05)
Observations	1,231	1,224	1,234
Baseline controls	Yes	Yes	Yes
P-value (Treated + Treated * Husband worked away = 0)	0.19	0.74	0.05
P-value (Spillover + Spillover * Husband worked away = 0)	0.72	0.40	0.61

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Husband worked away is a dichotomous variable equal to one if, at baseline, the husband worked in a different sector for at least four consecutive nights over the past three months.
The last two rows show the p-value of the treatment and spillover effects on women whose husband did work away.

Table A16: Impact on descriptive norms: Beliefs about IPV attitudes and prevalence in the community

Panel A: Women					
	(1)	(2)	(3)	(4)	(5)
	What share of men in the village do you believe...				
	Hold inequitable gender attitudes	Think wife should obey husband	Get jealous when wife speaks with other men	Kick/drag/beat up wife	Physically force wife to have sex with them
Treated couple	0.09**	-3.28**	0.91	3.88***	8.59***
	(0.04)	(1.42)	(1.43)	(1.20)	(1.60)
Spillover couple	0.03	-1.59	1.45	0.71	3.57**
	(0.04)	(1.40)	(1.53)	(1.26)	(1.54)
Observations	2,042	2,007	2,004	2,005	1,893
Baseline controls	Yes	Yes	Yes	Yes	Yes
Control mean	0.00	66.10	57.56	32.60	26.99
Treated vs Spillover p-value	0.17	0.28	0.76	0.02	0.01
Panel B: Men					
	What share of men in the village do you believe...				
	Hold inequitable gender attitudes	Think wife should obey husband	Get jealous when wife speaks with other men	Kick/drag/beat up wife	Physically force wife to have sex with them
Treated couple	0.10***	-3.49**	1.71	4.83***	8.56***
	(0.03)	(1.51)	(1.60)	(1.02)	(1.25)
Spillover couple	0.06*	-0.43	1.16	2.55***	2.19**
	(0.03)	(1.50)	(1.52)	(0.84)	(1.09)
Observations	2,042	1,985	1,982	1,986	1,847
Baseline controls	Yes	Yes	Yes	Yes	Yes
Control mean	0.00	54.93	45.98	18.76	14.05
Treated vs Spillover p-value	0.22	0.09	0.78	0.06	0.00

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, and differences between effects on trained and spillover couples, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Panel A shows the results for women's second-order belief regarding men in their village. Panel B shows the results for men's second-order belief regarding men in their village.
Column 1 uses a normalized composite index of beliefs about inequitable gender attitudes. Columns 2–5 use dependent variables expressed in percentage points.

Table A17: Heterogeneous treatment effects by husband's baseline mental well-being indicators

	(1)	(2)	(3)	(4)	(5)	(6)
	No. of husband's negative experiences during '94 conflict			Husband witnessed physical IPV against mother as a child		
	Emotional IPV	Physical IPV	Sexual IPV	Emotional IPV	Physical IPV	Sexual IPV
Treated couple	0.10	-0.02	0.09	0.09**	0.05	0.06
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.05)
Spillover couple	0.13**	0.12*	0.19***	0.06	0.17***	0.17***
	(0.06)	(0.07)	(0.05)	(0.04)	(0.05)	(0.04)
Treated * Characteristic	-0.01	0.02	0.00	-0.03	-0.01	0.04
	(0.02)	(0.02)	(0.02)	(0.06)	(0.06)	(0.07)
Spillover * Characteristic	-0.02	-0.00	-0.01	0.01	-0.07	-0.01
	(0.02)	(0.02)	(0.02)	(0.06)	(0.07)	(0.07)
Characteristic	0.00	-0.02	-0.00	-0.03	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.05)	(0.04)	(0.04)
Observations	1,972	1,958	1,970	1,366	1,354	1,360
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: IPV = intimate partner violence; No. = number; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
In columns 1–3, the treatment is interacted with the husband's number of negative experiences during the 1994 conflict.
In columns 4–6, the treatment is interacted with a dichotomous variable equal to one if the husband witnessed physical IPV against his mother as a child. All the interacted variables were asked at baseline.

Table A18: Additional specifications: First difference, PDS Lasso, and probit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First difference			PDS Lasso			Probit		
	Emotional IPV	Physical IPV	Sexual IPV	Emotional IPV	Physical IPV	Sexual IPV	Emotional IPV	Physical IPV	Sexual IPV
Treated couple	0.04	0.05*	0.11***	0.07***	0.04*	0.09***	0.21**	0.14**	0.28***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.08)	(0.07)	(0.07)
Spillover couple	0.06*	0.13***	0.20***	0.07***	0.11***	0.15***	0.19***	0.33***	0.46***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.07)	(0.07)	(0.07)
Observations	1,941	1,911	1,927	1,974	1,960	1,972	1,974	1,960	1,972
Baseline controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Note: IPV = intimate partner violence. PDS = post-double selection. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Columns 1–3 report results of the first-difference estimation. Columns 4–6 report results with additional control variables selected via PDS Lasso. Columns 7–9 report results of a probit estimation.

Table A19: Additional specifications: Instrumental variable treatment-on-the-treated

	(1)	(2)	(3)	(4)	(5)	(6)
	Emotional IPV		Physical IPV		Sexual IPV	
	First Stage	Second stage	First Stage	Second stage	First Stage	Second stage
Treated couple assignment	0.96***		0.96***		0.96***	
	(0.01)		(0.01)		(0.01)	
Spillover couple	0.04***	0.06***	0.04***	0.11***	0.04***	0.16***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Treated couple complier		0.07***		0.05**		0.10***
		(0.03)		(0.02)		(0.03)
Observations	1,974	1,974	1,960	1,960	1,972	1,972
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes

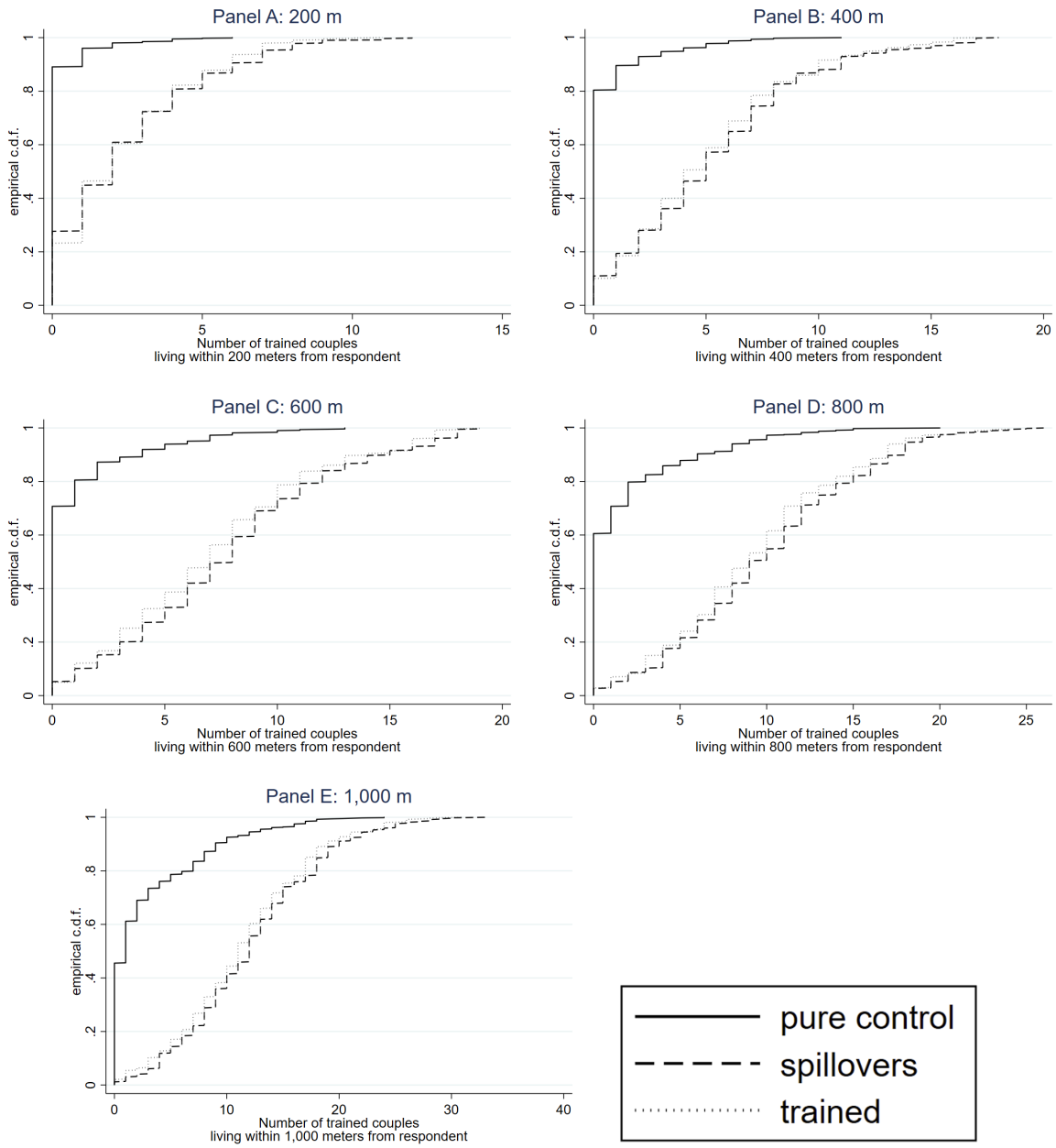
Note: IPV = intimate partner violence. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show two-Stage least squares (2-SLS) estimates of direct treatment effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing. Treatment assignment is used as instrument for actual treatment.

Table A20: Recoding missing refuse-to-answer violence questions as ”yes” and ”no”

	(1)	(2)	(3)	(4)	(5)	(6)
	Emotional IPV		Physical IPV		Sexual IPV	
VARIABLES	Refuse as "Yes"	Refuse as "No"	Refuse as "Yes"	Refuse as "No"	Refuse as "Yes"	Refuse as "No"
Treated couple	0.07**	0.08***	0.04*	0.06**	0.10***	0.10***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Spillover couple	0.06***	0.05**	0.12***	0.09***	0.16***	0.16***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Observations	2,008	1,995	1,999	1,999	2,008	2,008
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: IPV = intimate partner violence; OLS = ordinary least squares. Strata fixed effects included. Standard errors are clustered at village level, shown in rounded brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Results show OLS estimates of direct treatment and spillover effects, controlling for the baseline value of dependent variable (recoded to 0 if missing), as well as a dummy for whether the baseline response was missing.
Columns 1, 3, and 5 present the results in which all refuses-to-answer were recoded as "yes". In columns 2, 4, and 6, all refuses were recoded as "no".

Figure A1: How 'pure' are pure control couples? Number of treated couples living within various distances of couples.



Appendix B Additional Sampling Details

The sampling frame was chosen as follows. Within Eastern Province, and in partnership with province and district authorities, the Government of Rwanda selected four districts and eight sectors understood to have high rates of IPV and with a high concentration of VSLAs through which the intervention could be delivered. Within these eight sectors, existing CARE VSLAs were identified within 128 villages. From these 128 villages, the researchers ruled out 30 villages with fewer than three VSLAs, and villages that had VSLA meeting points that were very close to the VSLA meeting points of another village. This was done to increase the distance between control and treatment villages, minimizing the risk of contamination. There is no rule preventing residents of one village from participating in VSLAs based in another village. In the selected eight sectors and 98 remaining villages, VSLAs were then selected for inclusion in the sampling frame. In villages with three VSLAs exactly, all VSLAs were selected. In villages with more than three VSLAs, VSLAs in the village were ranked by the number of members. The three VSLAs with the greatest number of couples were selected for the sample to maximize the probability of having sufficient subscription to the intervention and to maximize its cost-effectiveness.

Appendix C Summary of Ethical Context and Protocols

Context

This study was conducted following several guidelines on the ethical research of IPV (Ellsberg et al., 2001; Heise and Hossain, 2017; WHO, 2016). We received ethics approval for the study from Oxford University and the Rwandan National Ethics Committee, and received a research visa from the National Institute of Statistics, Rwanda.

The Government of Rwanda partnered with the World Bank to identify and test a promising IPV prevention model — a gender transformative couples training program. If it was found to be effective, the program could form the basis of a scaled up program that could be delivered to many more couples in the country. The couples program tested here was based on a more intensive program that was tested and results have since been presented in two papers (Chatterji et al., 2020; Dunkle et al., 2020). The *Indashyikirwa* couples curriculum, 90% of which is shared with this couples training curriculum, was carefully adapted, extensively piloted, and revised to be appropriate for the context. The positive early anecdotal findings of the *Indashyikirwa* study, and the positive findings of a similar IPV prevention program for new parents in Rwanda (Doyle et al., 2018) — noting the significant differences in design and analysis between these studies and ours — suggested the program might work. Before scaling up a more cost-effective version of the program that condensed one element of the program — couples activism — into the couples training curriculum delivered to *all* trained couples rather than a subset, the government, supported by the World Bank, decided to evaluate it first and measure anticipated positive spillover effects of the program to others in the community (other VSLA members who were not treated). The results of this study were intended to inform a decision on scaling up the program. Similar couples-based IPV prevention programs are being delivered in other countries.

Study Protocols Summary

1. Consent
    - Easy to understand, read aloud to respondents to not disadvantage illiterate respondents.
    - Made clear that: there was no penalty for refusing to participate in the survey (they would still be eligible for the program); their data were private and no one except the survey company would know their name or any other identifying information; that identifying data were kept only to follow up with them for endline; and identifying data would be kept separate from their answers to the survey questions.
    - Asked whether they had been coerced into answering the survey or participating in the program.

- Consent forms had contact details of the research team and ethics committee in case anyone had follow-up questions.

2. Enumerators

   - Trained in conducting sensitive gender-based violence (GBV) surveys and understanding GBV causes.
   - Interviews held by enumerators of the same sex as respondents.
   - Almost all had previous experience conducting surveys related to GBV in Rwanda.
   - Never interviewed anyone they knew or in villages they were from.
   - Provided with support and debriefing from the GBV counselors assisting with data collection to minimize the risk of vicarious trauma.

3. Data collection

   - IPV questions asked using audio computer-assisted self-interviewing (ACASI), which increases privacy and made respondents feel more comfortable refusing to answer.
   - Made clear anyone could refuse any questions and end the interview at any time.
   - Men not asked about their perpetration of violence against women (asked about likelihood of becoming violent in a set of hypothetical situations).
   - Interviewed in private settings where they could not be overheard.
   - Consent form and introduction to the community as a survey about families and health, not about IPV or women.

4. Referrals and support

   - Counselors with experience supporting survivors of trauma and GBV were present at data collection sites during data collection. The counselors were available to speak with respondents who became distressed during the interview, and could follow up with women and men afterwards to help them access services they wanted.
   - Most survey sectors had women's safe spaces or One Stop Centers to support survivors of IPV with emergency shelter, medical, legal, and justice support services.
   - Respondents were offered small, private information sheets about the nearest support services — health, medical, and legal — for anyone including themselves or community members who might want help.

5. Data privacy

   - No organization (research team, World Bank, government) except the survey firm had any identifying data linking respondents to their answers. The encryption key was not shared outside a small team at the survey firm.
   - Answers were stored on an encrypted data collection platform. Uploaded daily and removed from password-protected tablets.

Structured Ethics Section

Following the approach of Asiedu et al. (2021), we summarize additional ethical considerations and features of this study below.

1. Equipoise

   Is there policy equipoise? That is, is there uncertainty regarding participants' net benefits from each arm of the study relative to the other arms and to the best possible policy to which participants could have access? If not, ethical randomization requires two conditions related to scarcity: (1) Was there scarcity, i.e., did the inclusion of multiple arms change the expected aggregate value of the programs delivered? (2) Do all ex ante identifiable participants have equal moral or legal claims to the scarce programs?

   This RCT does not violate policy equipoise because there is genuine uncertainty among experts that receiving the couples training program is better for the recipient than not receiving it. IPV prevention programs are of increasing interest to policy makers but there is no consensus on their effectiveness, with most existing causal studies of group-based IPV prevention programs finding they have null or small beneficial impacts (see introduction for a full discussion of the existing literature and our contribution to it). We expected beneficial effects of this program, but were not sure of the extent of the benefit or of its cost-effectiveness. Given the large share of women experiencing IPV worldwide and policy desire for programs to help tackle it, and an increasing proliferation of IPV prevention programs, we believe there is an ethical obligation to conduct rigorous research on their impacts, as challenging and sensitive as the topic is.

   The study has three groups: couples receiving the couples training program, spillover couples (untreated couples living in treated villages), and a control arm. The impact of each of these arms is uncertain, and similar interventions elsewhere have yielded both positive and null results. Each of the treatment arms and control arm were provided access to the government's rollout of safe spaces and training for community opinion leaders.

2. Role of researchers with respect to implementation

   Are researchers "active" researchers, i.e. did the researchers have direct decision-making power over whether and how to implement the program? If yes, what was the disclosure to participants and informed consent process for participation in the program? Providing IRB approval details may be sufficient but further clarification of any important issues should be discussed here. If no, i.e., implementation was separate, explain the separation.

The Government of Rwanda, together with CARE Rwanda and RWAMREC carried out the implementation, independent of the researchers, although the research team did the randomization. The study had no impact on the number of people who received the couples training during the study period, although it did determine who received the curriculum via the random assignment. The constraint on the number of training participants was determined ex ante by the program budget available.

The researchers were employed by the World Bank, which was the same organization providing funding to the government on IPV prevention-related activities, including the couples training program studied. However, the research team was part of a separate unit and was not in charge of the program implementation, funding, contracting, or monitoring, and did not have direct decision-making power over implementation decisions. Before implementation started, the research team provided technical assistance to the government by hiring expert consultants from the *Indashyikirwa* pilot study to reflect on lessons learned from that program's implementation and make some small changes to the couples training curriculum to reflect lessons learned. Modifications resulted in the condensing of the *Indashyikirwa* two-week stand-alone couples activist training into two extra sessions for all couples participating in the training.

Our role as researchers consisted of: suggesting that the government consider condensing the couples activist training to two sessions for all couples to test a possibly more scalable cost-effective IPV prevention program; funding the IPV training curriculum experts to reflect on earlier lessons learned and update the curriculum; within the intervention sectors selected by the government, randomly assigning villages to experimental arms; with the government, securing National Ethics Review Board approval and a Statistics Institute Research visa; securing University Ethics approval for human subjects research; funding the sensitization of study participants; based on sensitization public lottery results, randomly assigning couples to treatment, spillover, or control arms; designing and supervising the conducting of surveys of participants; analyzing the data; interpreting the results; and writing this paper. Researchers and their employer organization did not directly provide any of the interventions, or parts thereof, to participants, nor did they interact directly with participants.

3. Potential harms to participants or nonparticipants from the interventions or policies
   Does the intervention, policy, or product being studied pose potential harm to participants or non-participants? Related, are participants or likely affected non-participants particularly vulnerable? Also related, are participants' access to future services or policies changed because of participation in the study? If yes to any of the above, what is being done to mitigate such risks?

While no similar programs that the researchers are aware of have caused increases in IPV, there is evidence that some policy reforms or natural experiments that improve women's outside options or employment relative to men can either increase IPV or decrease IPV. Yet there is no clear understanding of the contexts in which improving women's situation increases or decreases the risk of IPV. There is some evidence from Rwanda that improving women's bargaining power reduces IPV (LaMattina, 2017; Sanin, 2023).

Therefore, this program did have the potential to pose a risk of increased IPV to treated women. Spillovers from social programs are typically small, so while it seemed unlikely the programs would cause spillover effects, it was also possible the program would have a negative effect on non-participants living in treated communities.

However, the Government of Rwanda's Ministry of Gender and Family Promotion wanted to reduce IPV in the country and this curriculum and this couples training program were promising based on evidence from similar programs in the Rwandan and other nearby country contexts (Chatterji et al., 2020; Doyle et al., 2018). To minimize risks, the curriculum upon which the intervention was based had been piloted and developed over several years in Rwanda, with intensive support and supervision from IPV prevention experts, and with the same NGOs implementing this program, who by the time of this study had been working on IPV prevention programs in Rwanda for over a decade.

The participants were recruited through the lead NGO's, CARE Rwanda, existing network of VSLAs in primarily rural parts of Rwanda's Eastern province (as selected by the government). VSLAs offer an opportunity for people without other financial mechanisms to save and borrow and are typically targeted at low-income households. Our comparison of the study sample and random community baseline sample suggests they are low income with high levels of food insecurity, though very similar to the average person living in these rural villages.

The researchers were not "active," and instead were observers of the program that would have been undertaken regardless of the research. Measuring the (expected) benefits or harms from the program was one reason for conducting it given the increasing use of similar IPV prevention programs around the world.

4. Potential harms to research participants from data collection (e.g., surveying, privacy, data management) or research protocols (e.g., random assignment)?
   Are data collection and/or research procedures adherent to privacy, confidentiality, risk-management, and informed consent protocols with regard to human subjects? Are they respectful of community norms, e.g., community consent, not merely individual consent, when appropriate? Are there potential harms to re-

search staff from conducting the data collection that are beyond "normal" risks? *Because these are all issues covered by most IRB processes, a sufficient explanation for a "yes" response may be to provide the IRB approval numbers for all IRBs that have approved the project. However, if there are particular issues that are important to discuss, please do so here.*

Related to these risks, approval was obtained from: Rwandan National Ethics Committee (722/RNEC/2018) and the Oxford University Central University Research Ethics Committee (R50347), and a research visa was granted by the National Institute of Statistics Rwanda (0676/2018/10/NISR). A summary of the steps taken is listed above.

These approvals and the above summary also discuss potential harms to research staff and mitigation strategies such as training and regular debriefs with study counselors and vicarious trauma training, and research tools to limit exposure to vicarious trauma.

7. Scarcity and counterfactual policy
   (Relevant for randomized controlled trials only): Did the inclusion of random assignment to treatment and/or control arms cause a change in the expected aggregate value of programs or products delivered?

   The program budget (independent of the study) limited the number of couples who could participate in the training program. Given this scarcity limiting scale, the random assignment merely influenced who got the program, not how many couples received it. Had the study not taken place, the same program would have been implemented anyway.

9. Researcher independence
   Were there any contractual limitations on the ability of the researchers to report the results of the study? If so, what were those restrictions, and who were they from?

   The researchers had unrestricted intellectual freedom to report the results of the study.

10. Financial conflicts of interest
    Do any of the researchers have financial conflicts of interest with regard to the results of the research?

    No.

11. Reputational conflicts of interest
    Do any of the researchers have potential reputational conflicts of interest?

No.

12. Feedback to participants or communities

    Is there a plan for providing feedback on research results to participants or communities? If yes, what is the plan? If not, why not?

    This was not yet common practice at the time this study was conducted. However, had it been, we would have been concerned about the challenge of successfully communicating the nuance of the research results given the sensitivities of the topic and the results. We would have been concerned that reporting the results of this study to participants could put them at increased risk of harm. However, we think that this issue is something that needs further discussion and guidance related to sharing negative results around sensitive topics with participants, noting that in theory, providing feedback on research results to participants respects their agency and is desirable if possible.

13. Foreseeable misuse of research results

    Is there a foreseeable and plausible risk that the results of the research will be misused and/or deliberately misinterpreted by interested parties to the detriment of other interested parties? If yes, please explain any efforts to mitigate such risk.

    No.

14. Other ethics issues to discuss

    Are there any other issues to discuss or issues that relate to a combination of the above? If so, please discuss here.

    No.

Appendix D Deviation from Pre-Analysis Plan

The specification and the variable definitions used in this paper were described in the pre-analysis plan (PAP) on the AEA Trial registry AEARCTR-0002282 prior to the research team receiving treatment-identifying endline data. The present section highlights the differences between this PAP and the final version of the paper, covering the pre-specified primary and secondary outcome analysis. Given the program caused the opposite treatment effects of those anticipated (an increase rather than decrease in IPV), and thus the opposite effects for which the exploratory mechanism analysis had been anticipated, we sometimes deviated from the exploratory mechanism analysis anticipated in the PAP and focused on exploratory analysis for possible mechanisms that could explain the observed increase in IPV. This means the paper includes additional outcomes and exploratory mechanism analysis to what was included in the PAP, and excludes some outcomes outlined in the PAP. We include results for all pre-specified secondary outcomes that are not in the paper in a supplementary appendix.

1. Primary outcomes : No deviation from the PAP
2. Secondary outcomes present in the PAP and absent from the paper (in supplementary appendix):
    - Economic empowerment: Women's and household consumption, productive loans, income, food security, and labor force participation; relative consumption of women's preferred items, women's body mass index
    - Agency - Spousal control: Female mobility, economic coercion
    - Relationship quality and cooperation: Cooperation decisions in incentivized game, knowledge of spouse's economic behavior
    - Well-being: Happiness and stress
3. Secondary outcomes present in the paper and absent from the PAP
    - Last disagreement escalated into violence (table 2)
    - Husband spends for himself money set aside for the household (table 2)
    - Number of situations in which the husband might become violent (table 5)
    - Divergence in attitudes between spouses (table 5)
    - Want more decision-making on household expenses and childcare (table 4)
    - Marriage formalization (table 3)
    - Women's reported injuries from IPV (table A7)
    - Men's reported IPV (table A7)
    - Individual IPV questions (table A10 and table A11)

- Refusal to answer questions (table A20)
- Indicators of stigma and empowerment to report (table A13)