Sex, Power, and Adolescence: Intimate Partner Violence and Sexual Behaviors*

Manisha Shah
University of California, Berkeley & NBER

Jennifer Seager
George Washington University

Joao Montalvao
World Bank

Markus Goldstein
World Bank

July 20, 2023

Abstract

Adolescents in Sub-Saharan Africa have some of the highest rates of intimate partner violence across the globe. This paper evaluates the impact of a randomized controlled trial that offers females a goal setting activity to improve their sexual and reproductive health outcomes and offers their male partners a soccer intervention, which educates and inspires young men to make better sexual and reproductive health choices. Both interventions reduce female reports of intimate partner violence. Impacts are larger among females who were already sexually active at baseline. We develop a model to understand the mechanisms at play. The soccer intervention improves male attitudes around violence and risky sexual behaviors. Females in the goal setting arm take more control of their sexual and reproductive health by exiting violent relationships. Both of these mechanisms drive reductions in IPV.

*We would like to thank participants at Stanford, Harvard, MIT, UC Berkeley, NYU, Columbia, U of Michigan, BREAD, NBER Health, UCSD, UT Austin, UCLA, UC Riverside, WADES, OU, Rice University, Tinbergen, Erasmus University, Georgetown, George Washington University, and the World Bank for their valuable comments and feedback. We thank Debraj Ray and Arun Chandrasekhar for many helpful conversations. Gabriela Rubio and Jiayin Zhai provided excellent research assistance. We thank the late SK Tariquzzaman for tirelessly leading the fieldwork. We gratefully acknowledge funding for this project from the Hewlett Foundation and the Africa Gender Innovation Lab (GIL) at the World Bank. Empirical analysis is pre-registered through the AEA RCT registry (RCT ID: AEARCTR-0001305). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. For questions or comments please contact Manisha Shah at ManishaShah@berkeley.edu.

1 Introduction

Intimate partner violence (IPV) is a global public health epidemic; nearly one in three women will experience some form of IPV in her lifetime (World Health Organization, 2021). In Tanzania, 32% of ever-partnered 15-19 year olds report ever experiencing IPV and 25% report experiencing IPV in the last 12 months (World Health Organization, 2021). In addition to the direct negative effects of violence on women's outcomes, social norms that perpetuate IPV and the resulting lack of bargaining power with sexual partners affect females' ability to make safe choices around sexual and reproductive health (SRH). IPV is also associated with risky sexual behavior, such as low rates of modern contraceptive use, multiple partnerships, and larger age gaps between partners (Melesse et al., 2020; Nkata, Teixeira and Barros, 2019; DHS, 2016).

We implement a randomized controlled trial (RCT) with female and male adolescents in Tanzania to change these power dynamics around adolescent relationships with the goal of improving female SRH outcomes related to violence and risky sexual behaviors.[1] Our interventions build on an ongoing adolescent empowerment program (Empowerment and Livelihoods for Adolescents (ELA)) delivered to females through a network of 149 clubs in three regions of rural Tanzania. For females, we randomize invitations to participate in a goal setting activity aimed at motivating the adoption of safe behaviors to improve their SRH outcomes. In randomly selected communities, the boyfriends of ELA participants are invited to participate in an intervention using an innovative sport-based pedagogy that employs soccer-specific activities, metaphors, and language to educate and inspire them. The curriculum focuses on reshaping males' attitudes and behaviors around masculinity, gender-based violence, and sexual relationships. We collect baseline data on all female ELA participants and their boyfriends and resurvey them two years later.

Intent-to-treat (ITT) estimates show that female experience of IPV decreases by 0.190 of a standard deviation as a result of the male soccer (*Boys*) intervention and by 0.248 of a standard deviation as a result of the female goal setting (*Goal*) intervention. Impacts are

[1] This research received ethical clearance in country through the Tanzania National Institute for Medical Research (NIMR) (protocol NIMR/HQ/R.8a/Vol. IX/2247) and from the University of California Los Angeles Institutional Review Board (protocol # 16-000125).

significantly larger for females who were already sexually active at baseline, highlighting greater efficacy of the interventions for those more vulnerable to IPV. We develop a game theoretic model of SRH and IPV to interpret the causal link between the interventions and IPV. In the model, males and females have preferences for risky sex, and violence emerges when their preferences conflict—namely, when males want it and females do not.[2] If she says no to risky sex, he may inflict violence to get her to acquiesce. But if he does that, she may exit the relationship.

The *Boys* treatment can reduce IPV either by decreasing his net payoff from violence and/or by decreasing his net payoff from risky sex. The *Goal* treatment induces the female to set improved SRH goals, increasing her disutility from risky sex. This means she will say no to risky sex more often. This can result in an increase or decrease in violence, depending on her relative costs of exit versus violence. In order for violence to decrease as a result of the *Goal* intervention, females must exit more often in response to violence. Otherwise, IPV will increase.

Our empirical results show that reductions in IPV from the *Boys* treatment are driven by an improvement in male attitudes around violence, as well as SRH, suggesting a role for both a decrease in the net benefit of violence and a decrease in the net benefit of risky sex. For the *Goal* treatment, we find increased partner churn, with females less likely to be with the same partner as at baseline, implying female exit as the mechanism for decreased IPV. Interestingly, boyfriends in the *Goal* arm appear to be of higher quality at endline.

This paper makes several important contributions. First, traditional programming has often ignored males in SRH education programs or service provision because they are not the primary beneficiaries of the services (Jewkes, Flood and Lang, 2015); however, because of gendered power dynamics, males may control decisions surrounding sexual behavior that impact SRH outcomes (Varga, 2003). Due to the design of this study, we can causally estimate whether treating males improves female outcomes.

[2] Empirical evidence shows males use violence to obtain risky sex (see Raj et al. (2007); Teitelman et al. (2011); Alleyne et al. (2011); Kalichman et al. (1998)), and this is supported by evidence in our data that female experience of IPV and male perpetration of IPV are associated with lower reported condom use (see Table A1).

Second, recent evidence suggests that targeting adolescents with interventions focused on changing attitudes toward gender norms and risky behaviors can be effective (Edmonds, Feigenberg and Leight, 2021; Dhar, Jain and Jayachandran, 2022). Since adolescents are at an age where they are establishing a course for future relationships and have more malleable attitudes (Steinberg, 2015; Sheehan et al., 2017), interventions may have larger and longer-term effects. However, due to the focus of SRH programming on married couples (e.g.,Doyle et al. (2018); Minnis et al. (2015); Dunkle et al. (2020)) and individual adults (e.g., Pronyk et al. (2006); Roy et al. (2019)), we still know relatively little about how to improve adolescent SRH outcomes in low-income settings (besides cash and school- or club-based programming).[3]

Third, the economics literature on the causal mechanisms behind IPV has focused exclusively on married couples, where exit costs are relatively high, and on the role of bargaining over household income and resources as a primary driver of IPV outcomes among women (e.g., Haushofer et al. (2019); Hidrobo, Peterman and Heise (2016); Angelucci (2008); Bobonis, González-Brenes and Castro (2013); Erten and Keskin (2018); Aizer and Dal Bó (2009); Aizer (2010)).[4] Our model expands beyond IPV as a bargaining response over monetary resources by focusing on partnership bargaining in another critical realm—sexual relations. In addition, given our focus on adolescents, exit costs might be lower.

Fourth, we contribute to the small causal literature on the impact of sports programming on adolescents (Beaman et al., 2021; Ditlmann and Samii, 2016) and to scant evidence on the role of goal setting in low-income settings. As far as we know, this is the first evaluation of the application of goal setting to SRH in any setting.[5] Lastly, this

[3]Financial incentives and education-based interventions have been shown to reduce teen pregnancy, early marriage, HIV/AIDS and IPV (e.g., Baird, McIntosh and Özler (2011); Handa et al. (2015); Bandiera et al. (2020); Buchmann et al. (2021); Duflo, Dupas and Kremer (2015); Jewkes et al. (2008); Gibbs et al. (2020)).

[4]While these models allow for changes in the value of the female's outside option to play a role in mitigating violence (Haushofer et al., 2019; Angelucci, 2008; Bobonis, González-Brenes and Castro, 2013), they largely abstract away from the possibility of female exit from the relationship due to high normative and real costs of marital dissolution (e.g.,Erten and Keskin (2018)).

[5]Setting goals has been found to increase self-control and decrease present-biased behavior (Hsiaw, 2013), improve worker performance and productivity (Goerg, 2015), improve student performance on tests, entrance exams, and homework (Clark et al., 2020), decrease energy consumption (Harding and Hsiaw, 2014), increase savings (Choi et al., 2006), etc.

study provides low-cost, scalable solutions for decreasing IPV among adolescents. Most previous causal evidence on decreasing violence involves cash transfers or provision of income (Baranov et al., 2021; Kerr-Wilson et al., 2020), and our interventions are 3-6 times lower cost depending on the comparison intervention.

2 Study Design

2.1 Setting

This study was implemented in three regions of Tanzania—Dodoma, Iringa, and Mbeya—in partnership with BRAC Maendeleo. Mbeya is the largest of the three regions in terms of population at 2.7 million people as of the 2012 census, with Dodoma having a population of 2.2 million and Iringa just under 1 million people (National Bureau of Statistics et al., 2012). The average population size of study communities is about 3,000, and these are rural areas.

These regions were selected due to the presence of 149 adolescent female clubs (Empowerment and Livelihoods for Adolescents (ELA)) that BRAC began operating in Tanzania in 2009. This program started in Bangladesh and is also implemented in Uganda, Sierra Leone, South Sudan, and Liberia. ELA is an education-based intervention designed to empower adolescent females by providing a safe social space, life-skills training, and support in adolescent development. Female adolescents and youth are invited to participate in ELA. Participation is voluntary but members are expected to attend five days per week from 3-6PM. Each club averages 20 members and has a mentor who runs the programs. Previous research in Tanzania finds that 25% of the eligible population participated in ELA clubs and finds no significant selection into clubs (Buehren et al., 2017). While the evidence on ELA from Uganda and Sierra Leone is mostly positive in terms of decreasing unintended teen pregnancy and early entry into marriage or cohabitation (Bandiera et al., 2020, 2019), Buehren et al. (2017) find no positive impacts of ELA in our setting of Tanzania.

The current study builds on top of the ELA club structure to evaluate, via an RCT, complementary interventions. Figure 1 illustrates the overall design of the RCT. Treatment status was assigned at the ELA club level and at the individual level, depending on the treatment. At the ELA club level, the 149 clubs were randomly allocated to three groups of equal size, stratified by region: two treatment arms and one control arm. The control arm (49 clubs) maintained the status quo of ELA clubs. The two treatments arms are (i) *Supply* (50 clubs), which provided access to free contraceptives, and (ii) *Boys* (50 clubs), which layers a soccer intervention for males in these communities. At the individual level, a sub-sample of females across all three study arms were randomly selected, stratified at the club level, to receive an invitation to participate in the *Goal* treatment, an individual goal setting activity.

This paper focuses on the evaluation of the *Boys* and *Goal* treatments. The evaluation of the *Supply* arm, which produces null results due to no uptake of contraceptives (see Table B1), is discussed in detail in Shah, Seager and Rubio (2022). Although it will not be discussed further in this current paper, we control for this study arm in all analyses.

2.2 Data Collection

Figure 1 presents the baseline sample distribution across study arms. We conducted a baseline census of members of all 149 ELA clubs from August to October 2016. Club leaders provided a complete list of active members. Females enrolled in school were considered active if they attended ELA meetings at least twice a week. Out of school females were considered active if they attended ELA meetings three times per week. The census identified a population of 3,419 active members across the 149 clubs, and all active members were selected for survey. The female baseline survey occurred from September to December 2016, 2–5 months before any interventions were implemented and resulted in a final sample of 3,178 females. Surveys were completed with 92.9% of the total number of females listed during the census. The discrepancy reflects changes in participation in ELA clubs rather than refusals to participate in survey.

We also collected data on the male partners of our female sample. During the baseline survey, females were asked to list males with whom they were friends, males to whom they were attracted, and males with whom they were currently or historically having sex. This list of males served as the sampling frame for the male survey sample. All of the males listed as sexual partners in *Boys* communities were selected for survey, and, in all other communties, a random sample of males were selected from the lists. The males' baseline survey took place from December 2016 to February 2017. In total 1,466 males were surveyed at baseline, split roughly evenly between communities assigned to the *Boys* intervention (787 males) and all other communties (679 males).

Prior to endline data collection, another census of ELA members was conducted during May 2018. Endline data collection took place between June and August 2018 for both males and females, six to eight months after the end of all interventions. Of the 3,178 females in our baseline sample, 2,591 were successfully tracked to the endline survey, an overall tracking rate of 81.5%. This tracking rate is similar across survey treatments (81% in the control arm, 85% in *Boys*, and 80% of females invited to *Goal*) and is in line with tracking rates of studies in similar contexts (Bandiera et al., 2020). We do not find evidence of differential attrition according to treatment status or our outcomes of interest. We discuss attrition in more detail in section 7.

Baseline and endline adolescent surveys collected information on the adolescent's household and about the adolescent's sexual behavior, SRH knowledge and attitudes, education and time use, health, and socio-emotional skills. STI and HIV testing was also conducted, but prevalence was unexpectedly low at baseline, around 1% for both, so this data is not used in analysis as we are underpowered.

2.3 Interventions and Takeup

Soccer Intervention. The *Boys* arm intervention was implemented by Grassroot Soccer (GRS), an organization focused on empowering adolescent males through the power of soccer, educating them on sexual and reproductive health topics, preventing HIV, and increasing uptake of health-promoting services among youth (ages 10-19).[6] The activity-based curriculum uses soccer language and analogies to start conversations around healthy and responsible behaviors and uses soccer drills and games to reinforce key messages.

The curriculum included 11 one-hour soccer practices on topics related to risk be-

[6] While this is the ideal age for the intervention, Grassroot Soccer treated a few males older than 19 for this study, as some of the boyfriends named by females in *Boys* treatment communities were older than 19.

haviors, HIV/AIDS prevention, and intimate partner violence and respecting females. Coaches are available after practice for an additional 15-30 minutes in case males want one-on-one meetings to discuss more private issues. Ten of the practices are on SRH issues and one is on malaria. Of the ten classes on SRH issues, several touch on issues directly related to IPV. For example, in the *Communicate* lesson (lesson two), males are expected to name at least one local service for victims of rape and violence. One key message of this lesson is "In life, we should all stand up for girls and women to protect them from abuse" (Grassroot Soccer, 2013). Similarly in lesson three, *Risky Partners*, the key message is about having sex with individuals your own age and not pressuring younger females to have sex. In lesson ten, *Red Card*, males are given scenarios worthy of a red card, such as bus drivers requiring sex from female passengers, older partners pressuring younger females to have sex, and gender-based violence. See Appendix Table C1 for more details on the curriculum for all sessions.

Grassroot Soccer began implementing sessions during the second half of February 2017, continuing through December 2017. In each region, five coaches each ran three rounds of programming, resulting in a total of 15 teams of approximately 25 males per region. The soccer intervention primarily targeted males within ELA club members' social and sexual networks; however, the ELA and GRS interventions were independent of one another. This resulted in about 300 males enrolling (35% of the male survey sample).[7] Because we had funding for 1,000 males to participate, Grassroot Soccer enrolled around 700 additional males from communities assigned to the *Boys* arm. We followed the standard GRS protocol for recruitment via schools and the community. Ultimately, 1,090 males completed the soccer curriculum in *Boys* communities.

Goal Setting. For the goal setting activity, facilitators asked selected females if they were willing to set a goal to remain healthy and stay STI/HIV free for the following

[7] Males who enrolled in Grassroot Soccer look similar to males who did not enroll in terms of household wealth, communication with parents, and age, but are 13.3 percentage points more likely to be enrolled in school, which is consistent with GRS's target population, and had larger households. There is also evidence that GRS was more easily able to contact older males, which may be indicative of phone access and ownership. See Table A2 for more detail.

year.[8] If they agreed, facilitators went through the S.M.A.R.T. process of setting Specific, Measurable, Achievable, Relevant, and Timely goals (Doran, 1981), which is often used in cognitive-behavioral therapy (CBT). Females were asked to identify and commit to up to three specific strategies to achieve the goal. This initial activity took about 90 minutes and was done one-on-one with a trained facilitator in August 2017. We invited 865 females, who were randomly selected from the baseline sample across all 149 clubs, to participate in this goal setting activity. Of the 865 females invited to participate, 789 participated (91%).[9] Of the 789 participants, 113 females (14.3%) set three strategies, 383 females (48.5%) set two strategies, and 293 females (37.1%) set only one strategy. Figure 3 highlights that the most commonly identified strategy was to use a condom, followed by abstinence and being faithful. Females also wrote about why this goal was important for their future and what obstacles they might face in following through with their specific strategies.

Four months later, in December 2017, facilitators checked in with the females to see if they were implementing the strategies they set and asked them about behavioral constraints they might be facing in meeting these goals. These meetings were also one-on-one and lasted about 60 minutes.

In Table A3 we investigate which characteristics are correlated with setting and achieving more strategies using data from the baseline survey. Females whose responses indicate depression set and achieved fewer strategies.[10] Consistent with the psychological concept of self-efficacy, females with higher general self-efficacy scores set and achieved more strategies.[11] Females from relatively poorer households (e.g., with earthen floors) set and achieved fewer strategies.

Figure 2 shows the timing of the interventions relative to data collection.

[8] Oettingwen and Gollwitzer (2010) argue that framing goals in terms of positive outcomes (rather than preventing negative outcomes) is more effective.

[9] Of the 76 females who did not participate, only two refused. The rest were either unavailable at the time of the intervention or had moved away from the study area.

[10] Depression is measured using the Patient Health Questionnaire-2 (PHQ-2), where a score of three or higher is indicative of depression. The PHQ-2 includes the first two items of the PHQ-9 (Kroenke, Spitzer and Williams, 2003).

[11] Self-Efficacy is measured using the General Self-Efficacy Scale developed by Schwarzer and Jerusalem (1995). A total self-efficacy score that ranges from 10-40 was calculated. We then standardized this score using the mean and standard deviation of the score among females in control communities.

3 Outcomes and Sample characteristics

3.1 Outcomes

The primary outcomes in this paper are related to intimate partner violence and sexual activity.[12]

Intimate partner violence. For females, intimate partner violence (IPV) is based on responses to three questions that capture her experience of violence with her most recent partner within the last two years. These are standard questions on IPV from the Tanzania Demographic and Health Surveys (DHS, 2016). Violence categories include physical (pushing, shaking, or throwing something at her), psychological (threatening to hurt or harm her or someone she cares about), and sexual (being physically forced to have sexual intercourse). Interviews were conducted in private and confidentiality was ensured. In cases where females reported violence, they were provided resources to seek support.

We generate indicators for psychological, physical, and sexual violence happening often. In addition, we generate the same indicators for violence occurring in the last year. We then generate an overall index across all six indicators. We standardize each indicator at baseline and endline separately around the mean and standard deviation of females in control communities who were not assigned to the *Goal* treatment and take the unweighted average across items (following Kling, Liebman and Katz (2007)).

Sexual Activity. For sexual activity, we focus on behaviors that may be mechanisms through which the interventions operate, such as gender attitudes around violence and SRH, risk perceptions around STIs, and changes in sexual partnerships (both quantity and quality). These outcomes are measured at both baseline and endline.

For each group of outcomes, we create an overall index, following the same procedure as for IPV.

[12]We present definitions for all registered primary outcomes in Appendix Table D1.

3.2 Sample Characteristics and Baseline Balance

Table 1 presents summary statistics of the primary outcomes and demographic characteristics at baseline. In the control group at baseline (column 1), females are 16.5 years of age, 25% of the sample is sexually active, and between 3.5% and 5.4% of females have experienced IPV in the past year, depending on the item.[13] In columns 3 and 4, we show the difference between the *Boys* treatment and the *Goal* treatment relative to the relevant control group, respectively. Overall, the RCT appears to be balanced across observed outcomes and demographics at baseline. In our main analysis, we focus on the balanced panel of 2,591 females who were surveyed at both baseline and endline. Table B2 presents baseline balance for this sub-sample. In addition, we show balance for the IPV outcomes for the sub-sample of females who were sexually active at baseline in Table B3.

We are interested in whether ELA participants are representative of adolescent females. To test for this, we compare our sample of ELA members to the random sample of females from the same communities in the baseline sample of Buehren et al. (2017) before ELA was introduced (see Table A4). We find no evidence of systematic differences. While ELA participants in Tanzania are less likely to have a child than non-participants, there is no evidence that they differ by education enrollment status, relationship status, engagement in income generating activities, or across several measures of household wealth (Buehren et al., 2017). Likewise, in Uganda, Bandiera et al. (2020) find little evidence of selection on observables into ELA participation.

4 Empirical Framework

We estimate intent-to-treat (ITT) impacts using difference-in-differences (DD), accounting for the cross-cutting randomization of the goal setting activity following Muralidharan, Romero and Wüthrich (2021). The specification is as follows:

[13]Rates of IPV measured in our data are consistent with estimates for equivalent populations from the Tanzania DHS (2016). See Table A5.

$$
\begin{aligned}
Y_{ict} = \alpha + \beta_1 \text{Boys}_c \times \text{Post}_t + \beta_2 \text{Goal}_i \times \text{Post}_t + \gamma_1 \text{Boys}_c \times \text{Post}_t \times \text{Goal}_i \\
+ \theta_1 \text{Goal}_i + \theta_2 \text{Post}_t + \theta_3 \text{Goal}_i \times \text{Boys}_c + X_{ict}^{'} \xi + \alpha_c + \epsilon_{ict}
\end{aligned}
\tag{1}
$$

where Y_{ict} is the outcome of interest for individual i in club c at time t, Boys_c and Goal_i are binary indicators for being assigned to the *Boys* and *Goal* treatments, respectively, and Post_t is a dummy variable that takes on the value one for the period after treatment is implemented. X_{ict} is a vector of controls including $\text{Supply}_c \times \text{Post}_t$, $\text{Supply}_c \times \text{Goal}_i$ and $\text{Supply}_c \times \text{Post}_t \times \text{Goal}_i$ to control for assignment to the *Supply* treatment as well as a set of individual characteristics. α_c is a vector of club fixed effects that control for club-level treatment assignment and to account for the stratification of the *Goal* treatment assignment. The standard errors ϵ_{ict} are clustered at the club level to account for the study design. The parameters of interest, β_1 and β_2, capture the ITT effects of the *Boys* treatment and *Goal* treatment, and γ_1 estimates the interaction between the two treatments.

The individual characteristics included in X_{ict} are age in years, highest grade attended, and binary indicators that the female never communicates with her mother about SRH topics and whether the female's household (i.e., parents) owns the house in which she lives, unless otherwise noted. We include these controls because they are strongly correlated with sexual activity and relationship status (Bruhn and McKenzie, 2009); however, the results are qualitatively similar if we do not include them (see Table A6). We estimate DD specifications because our primary outcomes, IPV and sexual activity, are relatively highly autocorrelated, which make them well-suited for DD analysis (McKenzie, 2012). We estimate sub-analyses by baseline sexual activity and partnership status.

5 Results

We present estimation results from equation 1 for IPV outcomes in Table 2. Columns 1 and 2 present the estimates for β_1 and β_2, and column 3 tests for equality of the treatment effects. Column 4 presents the outcome mean among the control group (females in ELA only communities who were not assigned to the *Goal* treatment) at endline and 5 presents the observations.

Table 2 shows that the *Boys* treatment reduces the IPV index 0.190 standard deviations (p=.022) compared to the control. Looking at the individual components of the indices, the *Boys* treatment reduces the various IPV outcomes between 1.1 and 3.7 percentage points. Table 2 also shows that the *Goal* treatment decreases the IPV index 0.248 standard deviations (p=.011) compared to the control. The individual components of the indices have magnitudes between 1.2 and 5.9 percentage points. Table A7 shows that females who are more engaged in the goal setting activity and set two to three strategies reap larger benefits than those who set only one or no strategies. Figures 4 and 5 present the ITT effects of the *Boys* treatment (β_1) and the *Goal* treatment (β_2) on the IPV index.

We cannot reject that the treatments effects are the same across arms (see column 3). Appendix Table A8 presents the coefficient estimate for γ and shows there are no additional reductions in IPV for females who were invited to goal setting in *Boys* treatment communities. This might be because each treatment alone reduces IPV prevalence to nearly zero.

The previously discussed impacts are based on the entire sample of females, starting at age 10, when almost no one is experiencing IPV. These outcomes become more salient as females age and become sexually active. At baseline, 25.7% of the sample reported ever having had sex. Figure 6 presents estimates of β_1 and β_2 from equation 1 for females who were and who were not sexually active at baseline separately. For this estimation, we re-center the IPV index at baseline and endline separately around females who were sexually active at baseline in control communities who were not assigned to the goal setting activity. Reductions in IPV are twice as large in the *Boys* arm (0.382 standard deviation reduction, p=.043) and 65% larger in the *Goal* arm (0.411 standard deviation reduction, p=.011) among females who were sexually active at baseline.[14] In the next section, we investigate potential mechanisms driving these reductions in IPV.

[14] We also look at heterogeneity by having a partner in the past two years at baseline and the results are consistent (see shown in Figure A1).

6 Conceptual Framework

Men may use violence against women when they disagree over sexual relations (Raj et al. (2007); Teitelman et al. (2011); Alleyne et al. (2011), Kalichman et al. (1998)).[15] The following model uses a simple game theoretic framework to explore the mechanisms driving violence during negotiations over sexual relations. We then explicitly discuss how the *Boys* and *Goal* interventions can change these interactions.

6.1 Model Setup

In the following one-shot, sequential game, nature first generates a male-female pair. Each player may derive positive or negative utility from engaging in risky sex, i.e. a trade-off between the pleasure of risky sex (e.g., unprotected sex) and its perceived expected cost (e.g., STI infection). When male and female preferences are aligned (i.e., either they both prefer risky sex or they both dislike it), there is no conflict and no chance of violence. We focus on the scenario where males gain positive utility from risky sex and females gain negative utility from risky sex. Formally, the payoff of the pair is $(s_m, -s_f)$, where $s_i > 0$, so that s_m denotes the male's net benefit from risky sex and s_f denotes the female's net cost from risky sex. We normalize the pair's payoff from being in a couple without risky sex to $(0, 0)$; thus, s_m and s_f are the additional benefits or costs associated with engaging in risky sex.

The game (depicted in Figure 7) is as follows.[16] First, the male decides whether or not to propose risky sex to the female. If he does not, the game ends with both players receiving a normalized payoff of zero. If he proposes, the female chooses yes or no. If she says yes, the game ends with the male receiving payoff $s_m > 0$ and the female receiving payoff $-s_f < 0$. If she says no, the male decides whether or not to respond with violence. If he chooses violence, the female can either (i) stay in the relationship, bearing the full cost of violence, resulting in payoffs $(v_m, -v_f)$, where v_m is a net benefit for the male and

[15]Table A1 also shows correlations between violence and risk behaviors in both the male and female data at baseline.

[16]Figure 7 denotes the complete information version of the game to give the reader a flavor of the game. However, we assume that payoffs are private information—while players know their own payoffs, they do not know each other's payoffs.

v_f is a net cost for the female;[17] or (ii) exit the relationship, incurring a cost of dissolving the relationship, resulting in payoffs $(-d_m, -d_f)$, where d_i is a cost for both males and females.[18] We assume throughout that $v_f > s_f$ for all s_f.

If $d_f < v_f$, then she exits the relationship when threatened with violence. We define these females as *high-types* (H). If $d_f > v_f$, the cost of exit is prohibitively high. We define these females as *low-types* (L). We let the cost of violence v_f be common across types and the cost of leaving d_f^k be type-specific, where $k = L, H$, such that $d_f^H < v_f < d_f^L$. The fraction of high-type females is given by α, with the remaining $1-\alpha$ being low-types. For both high-type and low-type females, s_f is distributed according to the cumulative distribution function $F(\cdot)$, which is continuous and strictly increasing everywhere.

The distribution of s_f and the values of α, d_f^H, d_f^L, and v_f are all common knowledge, but only females know their type. Likewise the distributions of s_m and v_m and the value of d_m are all common knowledge; however, the female does not know with certainty whether saying no will trigger a violent response from the male. We denote the probability of triggering violence by p.

If the female is high-type, she will say no if

$$s_f > pd_f^H. \tag{2}$$

If she is low-type, she will say no if

$$s_f > pv_f. \tag{3}$$

These conditions establish thresholds for female's cost from risky sex, above which a k-type female says no. Intuitively, a higher probability of a violent response, p, makes it increasingly difficult for both types to say no.

From conditions (2) and (3), we obtain two best response functions that map the probability that a high- and low-type female says no as a function of p:

[17] v_m can be positive or negative depending on the male's relative taste for and opportunity cost of violence.

[18] The female may still experience (some) violence at the point of leaving, but such violence is not chronic.

$$\rho^H(p) = 1 - F(pd_f^H), \tag{4}$$

and

$$\rho^L(p) = 1 - F(pv_f). \tag{5}$$

Conditional on the female saying no, the male, knowing α, ρ^H and ρ^L, uses Bayes' rule to calculate the probability that the female will exit if he responds with violence:

$$q(p) = \frac{\alpha\rho^H(p)}{\alpha\rho^H(p) + (1-\alpha)\rho^L(p)}. \tag{6}$$

Based on (6), the male chooses to respond with violence if

$$(1-q)v_m - qd_m > 0, \tag{7}$$

Rearranging, this condition can be written as

$$v_m > \frac{q}{1-q}d_m, \tag{8}$$

which establishes a threshold for the payoff from violence, v_m, above which the male responds with violence. Intuitively, the higher the probability q, the less attractive it is for him to respond with violence. Also, we show without loss of generality, all males with $s_m > 0$ will propose (see Appendix E.1 for the proof). It follows that the probability a male reacts violently to a no is given by

$$p(q) = 1 - M\left(\frac{q}{1-q}d_m\right), \tag{9}$$

where $M(\cdot)$ is the cumulative distribution function of v_m, which we assume to be continuous and strictly increasing everywhere. This gives the male's best-response function to q.

6.2 Equilibrium

An equilibrium is given by $\{q^*, p^*\}$ such that (i) ρ^H satisfies (4), ρ^L satisfies (5), and q^* satisfies (6), all evaluated at p^*; and (ii) p^* satisfies (9) evaluated at q^*. The function $p(q)$ is decreasing in q (i.e., males are less likely to respond with violence as females become more likely to exit). However, $q(p)$ can be increasing or decreasing in p. To ensure an increase in p results in an increase in $q(p)$, we impose a straightforward assumption: low-type females, who would suffer cost v_f, are more responsive to changes in p than high-type females, who have an exit option.

Proposition 1 *There exists a unique equilibrium* $\{q^*, p^*\}$.

A proof for Proposition 1 is provided in Appendix Section E.2. Figure 8 illustrates the equilibrium. The top two panels show the best response functions for low- and high-type females, respectively. The upward sloping curve in the bottom panel is the best response function $q(p)$. The downward sloping curve in the bottom panel is the best-response function $p(q)$. The curves $q(p)$ and $p(q)$ intersect once, showing there is a unique equilibrium.

The y-intercept for $q(p)$ is α because, as p goes to zero, all females will say no and q will converge to the share of high-type females in the population. Similarly, the x-intercept for $p(q)$, β, is the share of males for whom $v_m > 0$. This is because all males with $v_m > 0$ will respond with violence as q goes to zero.

6.3 Testable Predictions: *Boys* Arm

The Soccer curriculum in the *Boys* arm aims to reshape boys' attitudes towards IPV and teaches males the importance of avoiding risky behaviors to stop the spread of HIV/STIs. This has two implications from the model: the curriculum can decrease the net benefit of risky sex, s_m, and/or it can decrease the net benefit of violence, v_m.

If the former effect is strong enough to shift some males' s_m to be negative (i.e., he no longer wants risky sex), then this trivially decreases violence by decreasing the probability of a mismatch in preferences for risky sex (where $s_m > 0$ and $s_f < 0$), and, thus, of a potentially violent relationship.

A decrease in v_m unambiguously reduces p, the equilibrium probability the male responds with violence when the female says no (see Figure 9).

Proposition 2 *The Boys treatment unambiguously reduces violence.*

A proof of Proposition 2 is given in Appendix Section E.3. This decrease in violence results from a decrease in s_m and/or v_m.

6.4 Testable Predictions: *Goal* Arm

The *Goal* intervention strengthens females' commitment to adopt safe sexual behaviors to remain healthy. This translates to an increase in the net cost of risky sex, s_f, across the distribution of females, shifting $F(s_f)$ to the right and increasing $\rho^k(p)$ for all values of p. As a result, females will say no more often. In equilibrium, this can increase or decrease violence depending on whether the change in $\rho^k(p)$ is relatively larger for low-types vs. high-types. The intuition is that, if low-type females say no relatively more often, males are more likely to inflict violence as they learn the change is coming from low-types. Breakups become less likely and violence increases (see Figure 10(a)). If, on the other hand, high-type females say no relatively more often, males become less likely to inflict violence as females exit more often. Breakups become more likely and violence decreases (see Figure 10(b)).

Proposition 3 *The impact of the Goal intervention on violence is ambiguous.*

The necessary condition for the *Goal* intervention to decrease violence is a decrease in p^* and increase in q^*, such that breakups increase. A proof of Proposition 3 is given in Appendix Section E.4.

6.5 Empirical Evidence for Model Predictions

***Boys* Treatment.** In the model, the male's willingness to inflict violence is driven by both his preferences over risky sex (s_m) and the net payoff of violence (v_m). The *Boys* intervention could affect either of these channels. To empirically assess the explanatory power of these channels, we use the male survey data and estimate ITT impacts using DD on males' outcomes using the following specification:

$$\begin{aligned} Y_{ict} = \alpha + \beta_1 \text{Boys}_c \times \text{Post}_t + \beta_2 \text{Goal}_i \times \text{Post}_t + \gamma_1 \text{Boys}_c \times \text{Post}_t \times \text{Goal}_i \\ + \theta_0 \text{Boys}_c + \theta_1 \text{Goal}_i + \theta_2 \text{Post}_t + \theta_3 \text{Goal}_i \times \text{Boys}_c + X_{ict}^{'} \xi + \alpha_c + \epsilon_{ict}, \end{aligned} \quad (10)$$

where Y_{ict} is the outcome of interest for male i connected to a female in club c at time t, Boys_c is an indicator that the boy resides in a community assigned to the *Boys* intervention, Goal_i is an indicator that the female who is connected to the male was invited to the *Goal* treatment, and Post_t is an indicator for the post-treatment period. $X_{ict}^{'}$ is a vector of controls that includes individual characteristics equivalent to the controls for the females' models, except we control for whether the male speaks to his father about sexual reproductive health topics rather than his mother. Location fixed effects in α_c are at the region level to account for the level of stratification of treatment assignment to the *Boys* arm, and standard errors are clustered at the level of the female to whom the male is attached. For the males' estimates, the coefficient estimate of β_2 estimates the indirect treatment effect of his girlfriend being invited to the *Goal* intervention.

We present estimation results from equation 10 in Table 3. Columns 1 and 2 present the estimates for β_1 and β_2, and column 3 tests for equality of the treatment effects. Column 4 presents the outcome mean among the control group (males whose connected female is in an ELA only community who were not assigned to the *Goal* treatment) at endline and 5 presents the observations. Panel A captures v_m through males' attitudes toward violence and Panels B and C capture aspects of s_m through males/ risk perceptions of STIs and sexual activity.

Panel A of Table 3 focuses on violence attitudes, namely disagreement that "A woman should tolerate violence from her husband/partner," and agreement that "A man should not beat a woman under any circumstance." We code attitude responses so that higher values indicate improved attitudes. We see that the *Boys* treatment improves attitudes related to violence by 0.290 standard deviations (p=.016).[19] There is no comparable

[19]Table A9 shows that these shifts in attitudes are concentrated among males who were already sexually

impact of the *Goal* intervention on males' attitudes (Panel A, columns 2 and 3), which makes sense given it was males' girlfriends who were treated in this arm. The magnitude of the *Boys* impact is similar to RCT results from Dhar, Jain and Jayachandran (2022), who engaged adolescents in classroom discussions about gender equality. Table A10 shows that improvements in attitudes were larger among males who enrolled in Grassroot Soccer, with an improvement in the Violence attitudes index of 0.442 standard deviations (p=.002) among this group. This evidence is consistent with a decrease in the male's net payoff of violence (v_m) as a result of the *Boys* arm.

Panels B and C present evidence consistent with a decrease in males' net payoff of risky sex (s_m) as well. In Panel B, the *Boys* intervention increases males' perceptions around the likelihood of their friends having STIs. Previous research has found that increasing expectations of the likelihood of HIV infection reduces risky behavior and vice versa (e.g., Delavande and Kohler (2016)). Males in the *Boys* arm are 14 percentage points more likely to believe that a randomly selected female friend is very or somewhat likely to have an STI and 12.3 percentage points more likely to believe that at least 15 out of 100 randomly selected males his age in the community have an STI. In turn, males are 12.2 percentage points more likely to agree that girls have the right to demand condom use compared to males in control communities (p=.059). The Risk perception index shows an increase of 0.293 standard deviations for males in the *Boys* treatment arm ($p <$.000). In Table A10, we show that the treatment effect is larger among males who enrolled in Grassroot Soccer, with an increase of 0.370 standard deviations ($p <$.000) among this group. Again, as expected, there is no comparable impact of the *Goal* intervention on these outcomes (Panel B, columns 2 and 3).

A reduction in s_m implies fewer males proposing risky sex in the first place. While we do not have a direct measure of this, we try to capture it indirectly via sexual activity. Panel C of Table 3 presents impacts of the *Boys* arm on sexual activity as reported by males. Overall, males report a reduction in sexual activity of 0.098 standard deviation, driven by a reduction in currently having a partner by 6 percentage points and in the

active at baseline, precisely the group of males who would be perpetrating IPV and consistent with female reports of greater reductions of IPV among females who were already sexually active at baseline.

number of sexual partners by 0.116 fewer sexual partners on average (a 14% reduction). We corroborate these reports with female data in Panel A of Table 4. We find that females in *Boys* communities experience a 0.125 standard deviation (p=.032) reduction in sexual activity, primarily driven by a reduction in currently having a partner. In Table A10, we show that the reduction in the sexual activity index is larger among females for whom a male in their sexual network enrolled in Grassroot Soccer, showing a 0.297 standard deviation ($p <$.000) reduction in the sexual activity index among this group.

Figure 4 presents a summary of treatment effects across outcomes for the *Boys* treatment (β_1 from equation 10), highlighting improved violence attitudes and SRH risk perceptions for males and reductions in sexual activity. While we cannot identify the relative importance of v_m vs. s_m in reducing IPV, we note that both factors seem to be at play.

***Goal* Treatment.** Empirically we have observed an overall reduction in IPV among females assigned to the *Goal* arm. The model shows this can only be a result of an increase in the conditional probability of exit (q^*) and a decrease in the conditional probability of violence (p^*). An increase in q^* implies that females are more likely to exit relationships in response to violence.

We investigate relationship stability in Table 4. In Panel B, we estimate the treatment effects on the likelihood of being with the same partner as at baseline. As the outcome is a change from baseline to endline, we estimate a cross-sectional treatment-control model, controlling for the same baseline characteristics as in equation 1. The results in Panel B of Table 4, column 2, show evidence of increased relationship dissolution. Females invited to participate in the *Goal* treatment are 3.9 percentage points less likely to be with the same partner as at baseline than females in the control group (p=.070). However we cannot reject that this effect is the same as the *Boys* treatment, although the *Boys* treatment coefficient is not statistically significant. In Table A11, we restrict the sample to females who reported experiencing any IPV at baseline. Females in the *Goal* treatment are 30.5 percentage points less likely to be with the same partner as at baseline than females in the control group (p =.001) and this is significantly more likely in the *Goal* arm compared to the *Boys* arm (p =.003), recognizing sample sizes are small.

In addition, in Panel A of Table 4, females in the *Goal* treatment report having more total sexual partners ever, but are equally likely to be in a current partnership as the control group (both statistically different than the *Boys* treatment, see column 3), suggesting more partnership turnover in this arm. All of these results are consistent with an increase in q driven by the *Goal* treatment. Figure 5 presents a summary of treatment effects across outcomes for the *Goal* treatment arm (β_2), highlighting increased exits.

The model is built around the notion that females will exit violent relationships conditional on her opportunity cost of leaving being sufficiently low (i.e., that she is a high-type, $d_f^H < v_f$). Table A12 compares characteristics of females who exited relationships to those who did not, conditional on naming a sexual partner at baseline. Panel A shows that females who exited relationships are nearly twice as likely to be currently enrolled in school and are 13 percentage points less likely to be currently married or cohabiting, both of which are consistent with high-type characteristics. Panel B broadly suggests that females who exited are less likely to be experiencing IPV at endline, consistent with high-type females leaving violent relationships.

Even though our model does not speak to matching or dynamics over time, we now use our rich data to explore two potential consequences of increased break-ups with violent partners: quality of subsequent partners and displacement of violence to other females in the community. In Panel C of Table 4, we restrict the sample to females who report having partners at baseline and/or endline and utilize data from female reports of her boyfriends' characteristics. For each female, we average the characteristics of her boyfriends for age and school enrollment, as these characteristics are correlated with risky sex and partnerships (Agüero and Bharadwaj, 2014; Schaefer et al., 2017; Beauclair, Dushoff and Delva, 2018). We find that goal setting significantly increases average boyfriend quality by 0.265 standard deviations (p=.023). Next, we compare IPV outcomes of females invited to the *Goal* treatment to control group females. If violent partners are being displaced, we would expect a reported increase in IPV among control females that offsets the decrease in IPV among those in the *Goal* treatment. Table A13 shows that while IPV significantly decreases for females in the *Goal* arm, there is no offsetting increase among control females.

Treatment Interaction. Encouraging women to say "no" more often can have a backlash effect for low-type females, as demonstrated in Figure 10(a). In that world, additionally implementing the *Boys* intervention could mitigate this backlash. Table A8 presents the same results as Table 2 but includes the coefficient estimate for γ_1 from equation 1. While imprecisely estimated, the coefficients on the interaction between the two interventions are positive for the IPV index, indicating that, in this case, the interventions may substitute each other in terms of reducing IPV. As mentioned previously, this could be because each treatment alone reduces IPV prevalence to nearly zero.

7 Attrition

Tables 5 and 6 present analysis of sample attrition using baseline data for the females and males, respectively, to test whether attrition varies by treatment status and/or baseline characteristics. The outcome in all panels is an indicator equal to 1 if the female (male) attrited by endline. We find no evidence of differential attrition by treatment status overall or by baseline characteristics.

We show baseline outcome means by attrition status in appendix figure A2, which shows no evidence of differential attrition. There is some evidence that females in the *Goal* treatment who experienced physical abuse at baseline are less likely to attrit. If anything, this would imply positive bias in our estimate of the impact of goal setting on physical abuse, i.e., our impacts are a lower bound.

8 Cost-effectiveness

We now present evidence on the cost-effectiveness of our interventions. Given the lack of experimental studies that provide evidence on reducing IPV among adolescents in LMICs that include cost data, we benchmark our IPV impacts and costs against two studies that estimate the impact of cash transfer programs on IPV among married women in Kenya (Haushofer et al., 2019) and Ecuador (Hidrobo, Peterman and Heise, 2016). We acknowledge that cash transfer programs are designed to shift many other outcomes unrelated to violence and these comparisons should be considered with this in mind.

Table A14 summarizes our cost effectiveness comparison. The per-female cost of the *Boys* treatment is $41 and the per-female cost of the *Goal* treatment is $38. To ease comparison across treatments and studies, we normalize the cost of each treatment to a 0.25 standard deviation reduction in IPV—$54 for the *Boys* treatment and $38 for the *Goal* treatment per 0.25 standard deviation decrease.

Haushofer et al. (2019) find that $496 cash transfers to adult women in Kenya reduced physical violence by 0.26 standard deviation and sexual violence by 0.22 standard deviation. Transfers of equal value to their husbands reduced physical violence by 0.18 standard deviations. These imply a cost of $477 to $539 per 0.25 standard deviation reduction in IPV from cash given to women, and a $689 per 0.25 standard deviation reduction in IPV from cash given to their husbands. Similarly, Hidrobo, Peterman and Heise (2016) find that monthly transfers (cash or in-kind) of $40 to adult women in Peru over a six-month period, for a total of $240 per woman, reduces physical or sexual violence by 6 percentage points. This translates to a cost of $400 per 0.25 standard deviation reduction in IPV. This basic costing analysis suggests that our interventions are highly cost-effective in reducing IPV relative to cash and in-kind transfers.

9 Discussion and Conclusion

This paper presents causal evidence from a multi-level cluster and individual RCT and finds that offering males a soccer-based health intervention reduces female experience of IPV by 0.190 standard deviations on average. Similarly, offering females a goal setting activity reduces experience of IPV by 0.248 standard deviations. Reductions in IPV in both treatment arms are larger for females who were already sexually active at baseline.

Sexual relations can be the outcome of power relations between females and males. We evaluate interventions that each shift one side of the relationship. We develop a simple model to illuminate mechanisms behind the power relations that drive SRH and IPV outcomes. In our model, male decision-making around IPV is driven by his net payoffs from risky sex and violence and expectations around whether his partner will exit in response to proposals of risky sex and violence. Females decide whether to engage in risky sex based on their net payoffs from risky sex and the costs they face from violence relative to exiting the relationship, along with their expectations about the likelihood their partner will perpetrate IPV.

This model speaks directly to our interventions, which separately target adolescent males and females to shift the dynamics that allow for IPV at this critical juncture in male and female development. The reduction in IPV in the *Boys* arm is driven by an improvement in male attitudes toward IPV and risky sex. Mapping back to our model, this implies a reduction in the likelihood the male responds with violence when the female says no and a reduction in the likelihood the male proposes risky sex in the first place. On the other side, the *Goal* arm helps females set concrete strategies on how to improve their sexual and reproductive health, increasing the cost of risky sex. The model suggests this will lead to females saying no more often, which could decrease or increase violence depending on whether females can leave these relationships once threatened with violence. Our data suggests that more females are able to exit relationships when faced with violence, resulting in an overall reduction in violence.

While programming focusing on adolescents is increasing, there is still little evidence on what works to reduce IPV for this age group. These results provide evidence of two effective, inexpensive, and scalable interventions to reduce IPV experienced by adolescent females. Changing gender relations at this early stage of adulthood could potentially shift the life trajectory of young men and women, which is a fruitful avenue for future research. In addition, research to understand how these interventions work, together or separately, in higher-violence settings could provide important guidance on when and where to scale-up.

References

Agüero, Jorge M and Prashant Bharadwaj. 2014. "Do the More Educated Know More About Health? Evidence from Schooling and HIV Knowledge in Zimbabwe." *Economic Development and Cultural Change* 62(3):489–517.

Aizer, Anna. 2010. "The Gender Wage Gap and Domestic Violence." *American Economic Review* 100(4):1847–59.

Aizer, Anna and Pedro Dal Bó. 2009. "Love, hate and murder: Commitment devices in violent relationships." *Journal of Public Economics* 93(3):412–428.
URL: *https://www.sciencedirect.com/science/article/pii/S0047272708001370*

Alleyne, Binta, Victoria H Coleman-Cowger, Laurel Crown, Maya A Gibbons and Linda Nicole Vines. 2011. "The Effects of Dating Violence, Substance Use and Risky Sexual Behavior Among a Diverse Sample of Illinois Youth." *Journal of Adolescence* 34(1):11–18.

Angelucci, Manuela. 2008. "Love on the Rocks: Domestic Violence and Alcohol Abuse in Rural Mexico." *The BE Journal of Economic Analysis & Policy* 8(1).

Baird, Sarah, Craig McIntosh and Berk Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *The Quarterly Journal of Economics* 126(4):1709–1753.
URL: *https://doi.org/10.1093/qje/qjr032*

Bandiera, Oriana, Niklas Buehren, Markus Goldstein, Imran Rasul and Andrea Smurra. 2019. "The Economic Lives of Young Women in the Time of Ebola: Leassons from an Empowerment Program." World Bank Policy Research Working Paper 8760.

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

Baranov, Victoria, Lisa Cameron, Diana Contreras Suarez and Claire Thibout. 2021. "Theoretical Underpinnings and Meta-analysis of the Effects of Cash Transfers on Intimate Partner Violence in Low-and Middle-Income Countries." *The Journal of Development Studies* 57(1):1–25.

Beaman, Lori, Sylvan Herskowitz, Niall Keleher and Jeremy Magruder. 2021. "Stay in the Game: A Randomized Controlled Trial of a Sports and Life Skills Program for Vulnerable Youth in Liberia." *Economic Development and Cultural Change* 70(1):129–158.

Beauclair, Roxanne, Jonathan Dushoff and Wim Delva. 2018. "Partner Age Differences and Associated Sexual Risk Behaviours Among Adolescent Girls and Young Women in a Cash Transfer Programme for Schooling in Malawi." *BMC public health* 18(1):1–12.

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

Bruhn, Miriam and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics* 1(4):200–232.
URL: *https://www.aeaweb.org/articles?id=10.1257/app.1.4.200*

Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Naznee and Xiao Yu Wang. 2021. "A Signal to End Child Marriage: Theory and Experimental Evidence from Bangladesh.".

Buehren, Niklas, Markus Goldstein, Selum Gulesci, Munshi Sulaiman and Venus Lam. 2017. "Evaluation of an Adolescent Development Program for Girls in Tanzania." Policy Research Working Paper.

Choi, James J., David Laibson, Brigitte C. Madrian and Andrew Metrick. 2006. Saving for Retirement on the Path of Least Resistance. In *Behavioral Public Finance*. Russell Sage Foundation pp. 304–352.
URL: *http://www.jstor.org/stable/10.7758/9781610443852.14*

Clark, Damon, David Gill, Victoria Prowse and Mark Rush. 2020. "Using Goals to Motivate College Students: Theory and Evidence from Field Experiments." *The Review of Economics and Statistics* 102(4):648–663.

Delavande, Adeline and Hans-Peter Kohler. 2016. "HIV/AIDS-related Expectations and Risky Sexual Behaviour in Malawi." *The Review of Economic Studies* 83(1):118–164.
URL: *https://doi.org/10.1093/restud/rdv028*

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

DHS. 2016. Tanzania Demographic and Health Survey (DHS) and Malaria Indicator Survey (MIS) 2015-2016. Technical report Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) and Tanzania Mainland, Ministry of Health (MoH) and Zanzibar, National Bureau of Statistics (NBS) and Tanzania, Office of Chief Government Statistician (OCGS) and Zanzibar, ICF.
URL: *http://dhsprogram.com/pubs/pdf/FR321/FR321.pdf*

Ditlmann, R. K. and C. Samii. 2016. "Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel." *Journal of Peace Psychology* 22(4):380–392.

Doran, George T. 1981. "There's a S.M.A.R.T. way to write managements's goals and objectives." *Management Review* 70(11):35–36.

Doyle, Kate, Ruti G. Levtov, Gary Barker, Gautam G. Bastian, Jeffrey B. Blingenheimer, Shamsi Kazimbaya, Ancient Nzabonimpa, Julie Pulerwitz, Felix Sayinzoga, Vandana Sharma and Dominick Shattuck. 2018. "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* 13(4).
URL: *https://doi.org/10.1371/journal.pone.0192756*

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105(9):2757–97.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20121607*

Dunkle, Kristin, Erin Stern, Sangeeta Chatterji and Lori Heise. 2020. "Effective prevention of intimate partner violence through couples training: a randomised controlled trial of Indashyikirwa in Rwanda." *BMJ Global Health* 5(12).
URL: *https://gh.bmj.com/content/5/12/e002439*

Edmonds, Eric V, Benjamin Feigenberg and Jessica Leight. 2021. "Advancing the Agency of Adolescent Girls." *Review of Economics and Statistics* 1–46.

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

Gibbs, Andrew, Laura Washington, Nada Abdelatif, Esnat Chirwa, Samantha Willan, Nwabisa Shai, Yandisa Sikweyiya, Smanga Mkhwanazi, Nolwazi Ntini and Rachel Jewkes. 2020. "Stepping Stones and Creating Futures Intervention to Prevent Intimate Partner Violence Among Young People: Cluster Randomized Controlled Trial." *Journal of Adolescent Health* 66(3):323–335.
URL: *https://www.sciencedirect.com/science/article/pii/S1054139X19304847*

Goerg, Sebastian J. 2015. "Goal setting and worker motivation." *IZA World of Labor* 178:1–10.
URL: *10.15185/izawol.178*

Grassroot Soccer. 2013. Zinduka Coach's Guide. Technical report.

Handa, Sudhansu, Amber Peterman, Carolyn Huang, Carolyn Halpern, Audrey Pettifor and Harsha Thirumurthy. 2015. "Impact of the Kenya Cash Transfer for Orphans and Vulnerable Children on early pregnancy and marriage of adolescent girls." *Social science & medicine* 141:36–45.
URL: *https://www.ncbi.nlm.nih.gov/pubmed/26246032*

Harding, Matthew and Alice Hsiaw. 2014. "Goal Setting and Energy Conservation." *Journal of Economic Behavior and Organization* 107:209–227.

Haushofer, Johannes, Charlotte Ringdal, Jeremy P Shapiro and Xiao Yu Wang. 2019. "Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya." NBER Working Paper, No. 25627.

Hidrobo, Melissa, Amber Peterman and Lori Heise. 2016. "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador." *American Economic Journal: Applied Economics* 8(3):284–303.

Hsiaw, Alice. 2013. "Goal-setting and self-control." *Journal of Economic Theory* 148(2):601–626.
URL: *https://EconPapers.repec.org/RePEc:eee:jetheo:v:148:y:2013:i:2:p:601-626*

Jewkes, Rachel, M Nduna, J Levin, N Jama, K Dunkle, A Puren and N Duvvury. 2008. "Impact of Stepping Stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: cluster randomised controlled trial." *BMJ* 337.
URL: *https://www.bmj.com/content/337/bmj.a506*

Jewkes, Rachel, Michael Flood and James Lang. 2015. "From work with men and boys to changes of social norms and reduction of inequities in gender relations: a conceptual shift in prevention of violence against women and girls." *The Lancet* 385(9977):1580–1589.
URL: *https://www.sciencedirect.com/science/article/pii/S0140673614616834*

Kalichman, Seth C, Ernestine A Williams, Charsey Cherry, Lisa Belcher and Dena Nachimson. 1998. "Sexual Coercion, Domestic Violence, and Negotiating Condom Use Among Low-income African American Women." *Journal of Women's Health* 7(3):371–378.

Kerr-Wilson, Alice, Andrew Gibbs, Erika McAslan Fraser, Leanne Ramsoomar, Anna Parke, Hussaind MA Khuwaja and Rachel Jewkes. 2020. A rigorous global evidence review of interventions to prevent violence against women and girls. Technical report What Works to Prevent Violence Against Women and Girls Global Programme, Pretoria, South Africa.

Kling, Jeffrey R, Jeffrey B. Liebman and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1):83–119.

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

Kroenke, Kurt, Robert L. Spitzer and Janet B. W. Williams. 2003. "The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener." *Medical Care* 41(11):1284–1292.
URL: *http://www.jstor.org/stable/3768417*

Locke, Ediwn A and Gary P Latham. 1990. *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice-Hall.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99(2):210–221.
URL: *https://www.sciencedirect.com/science/article/pii/S030438781200003X*

Melesse, Dessalegn Y, Martin K Mutua, Allysha Choudhury, Yohannes D Wado, Cheikh M Faye, Sarah Neal and Ties Boerma. 2020. "Adolescent sexual and reproductive health in sub-Saharan Africa: who is left behind?" *BMJ Global Health* 5(1).
URL: *https://gh.bmj.com/content/5/1/e002231*

Minnis, Alexandra M, Irene A Doherty, Tracy L Kline, William A Zule, Bronwyn Myers, Tara Carney and Wendee M Wechsberg. 2015. "Relationship power, communication, and violence among couples: results of a cluster-randomized HIV prevention study in a South African township." *International journal of women's health* 7:517–525.

Muralidharan, Karthik, Mauricio Romero and Kaspar Wüthrich. 2021. "Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments." NBER Working Paper, No. 26562.

National Bureau of Statistics, Ministry of Finance, Finance Office of Chief Government Statistician President's Office, Economy and Development Planning. 2012. "Tanzania Population and Housing Census.".

Nkata, Hamida, Raquel Teixeira and Henrique Barros. 2019. "A scoping review on sexual and reproductive health behaviors among Tanzanian adolescents." *Public Health Reviews* (4).
URL: *https://doi.org/10.1186/s40985-019-0114-2*

Oettingwen, Gabriele and Peter M. Gollwitzer. 2010. Strategies of Setting and Implementing Goals: Mental Contrasting and Implementing Intentions. In *Social Psychological Foundations of Clinical Psychology*, ed. James E. Maddux and June Price Tangney. Guilford Press.

Pronyk, Paul M, James R Hargreaves, Julia C Kim, Linda A Morison, Godfrey Phetla, Charlotte Watts, Joanna Busza and John DH Porter. 2006. "Effect of a structural intervention for the prevention of intimate-partner violence and HIV in rural South Africa: a cluster randomised trial." *The Lancet* 368(9551):1973–1983.
URL: *https://www.sciencedirect.com/science/article/pii/S0140673606697444*

Raj, Anita, Elizabeth Reed, Elizabeth Miller, Michele R Decker, Emily Faith Rothman and Jay G Silverman. 2007. "Contexts of Condom Use and Non-condom Use among Young Adolescent Male Perpetrators of Dating Violence." *AIDS care* 19(8):970–973.

Roy, Shalini, Melissa Hidrobo, John Hoddinott and Akhter Ahmed. 2019. "Transfers, Behavior Change Communication, and Intimate Partner Violence: Postprogram Evidence from Rural Bangladesh." *The Review of Economics and Statistics* 101(5):865–877.
URL: *https://doi.org/10.1162/rest_a_00791*

Schaefer, Robin, Simon Gregson, Jeffrey W Eaton, Owen Mugurungi, Rebecca Rhead, Albert Takaruza, Rufurwokuda Maswera and Constance Nyamukapa. 2017. "Age-disparate Relationships and HIV Incidence in Adolescent Girls and Young Women: Evidence from Zimbabwe." *AIDS (London, England)* 31(10):1461.

Schwarzer, Ralf and Matthias Jerusalem. 1995. Generalized Self-Efficacy scale. In *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*. Windsor, UK: NFER-NELSON pp. 35–37.

Shah, Manisha, Jennifer Seager and Gabriela Rubio. 2022. "Supplying Contraceptives to Adolescents: When Pure Access does not Breed Use." Working Paper.

Sheehan, Peter, Kim Sweeny, Bruce Rasmussen, Annababette Wils, Howard S Friedman, Jacqueline Mahon, George C Patton, Susan M Sawyer, Eric Howard, John Symons, Karin Stenberg, Satvika Shalasani, Neelam Maharaj, Nicola Reavley, Hui Shi, Masha Fridman, Alison Welsh, Emika Nsofor and Laura Laski. 2017. "Building the foundations for sustainable development: a case for global investment in the capabilities of adolescents." *The Lancet* 390(10104):1792–1806.

Steinberg, Laurence. 2015. *Age of Opportunity: Lessons from the New Science of Adolescence*. New York: Eamon Dolan/Mariner Books.

Teitelman, Anne M, Julie Tennille, Julia M Bohinski, Loretta S Jemmott and John B Jemmott III. 2011. "Unwanted Unprotected Sex: Condom Coercion by Male Partners and Self-silencing of Condom Negotiation Among Adolescent Girls." *Advances in Nursing Science* 34(3):243–259.

Varga, Christine A. 2003. "How Gender Roles Influence Sexual and Reproductive Health Among South African Adolescents." *Studies in Family Planning* 34(3):160–172.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1728-4465.2003.00160.x*

World Health Organization. 2021. *Violence against women prevalence estimates, 2018: global, regional and national prevalence estimates for intimate partner violence against women and global and regional prevalence estimates for non-partner sexual violence against women*. World Health Organization: Geneva.

Figures and Tables

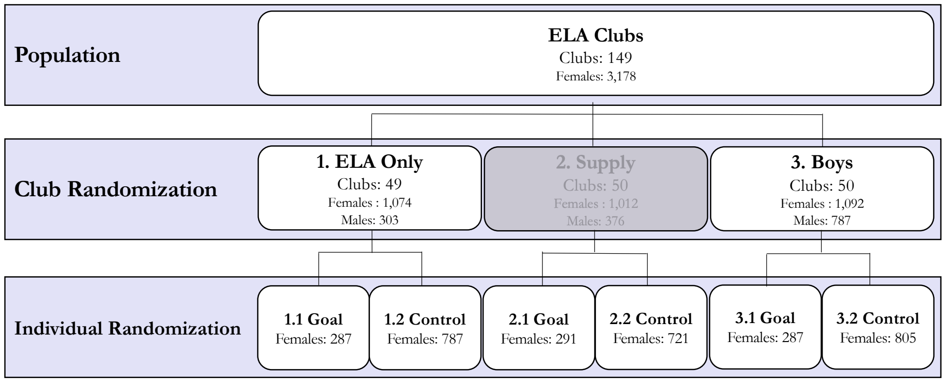


Figure 1 Study Design

Notes. This figure presents the overall study design. The study population, presented in the top box, is female participants in 149 ELA clubs at baseline in 2016. The middle box shows community-level randomization and the number of males and females surveyed at baseline in each community-level treatment arm. The bottom box shows the cross-cutting, individual-level *Goal* treatment.

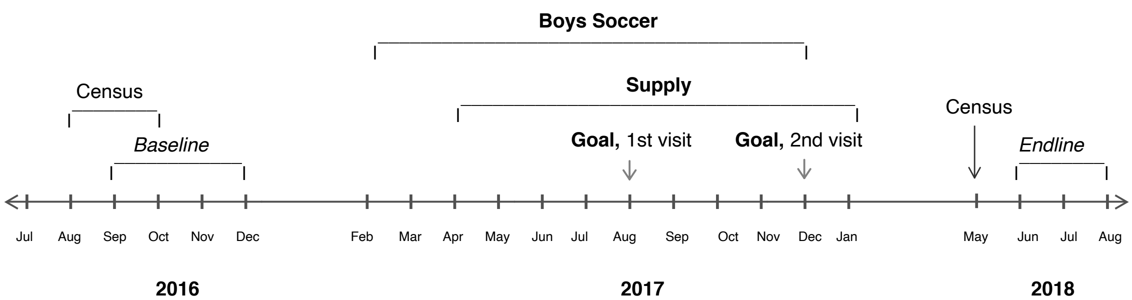


Figure 2 Study Timeline

Notes. This figure presents the study timeline.

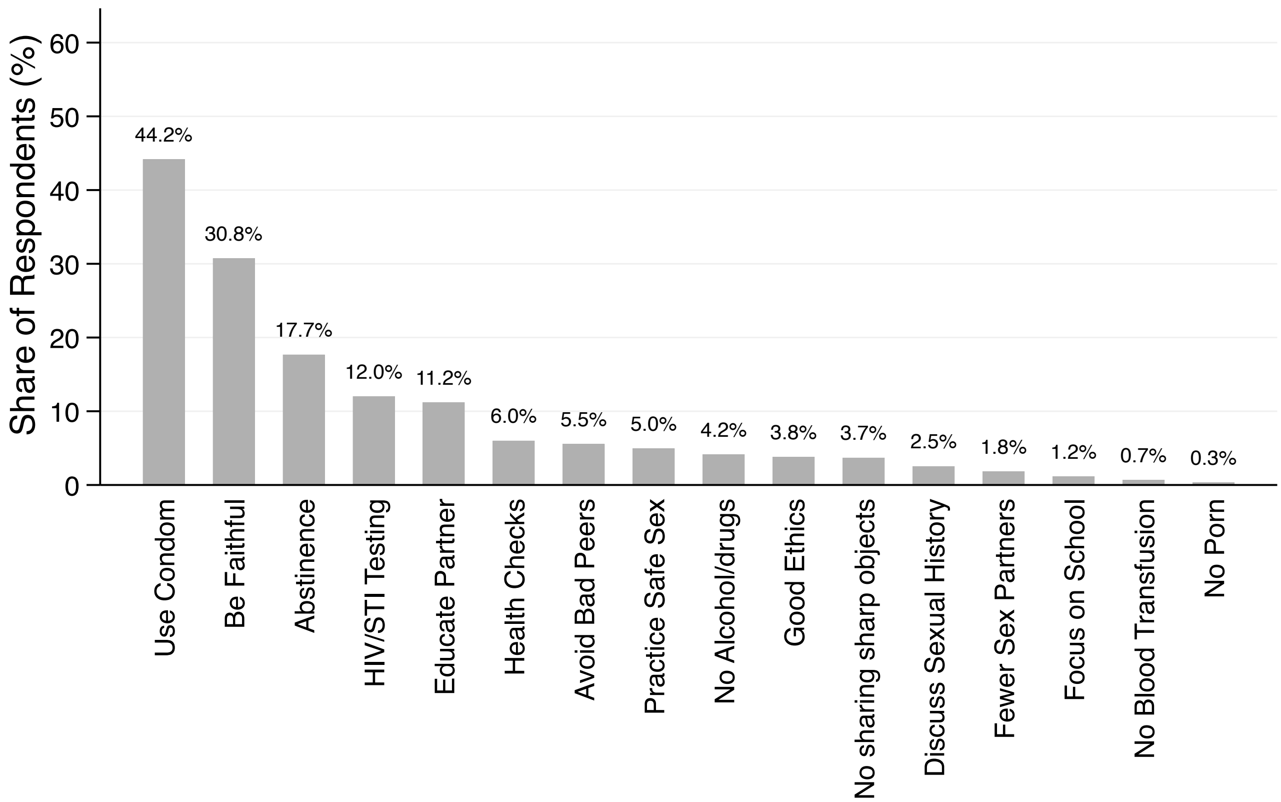


Figure 3 Strategies from Goal Setting Activity

Notes. This figure summarizes strategies identified during the goal setting activity. Each female was asked to identify 1–3 strategies. These strategies were categorized into 16 over-arching categories. The percent of females who set a strategy that fits in each category is presented above the bar. As females could set up to 3 strategies, the percentages above the bars do not sum to 100%.
Source. Female goal setting participants, first visit.

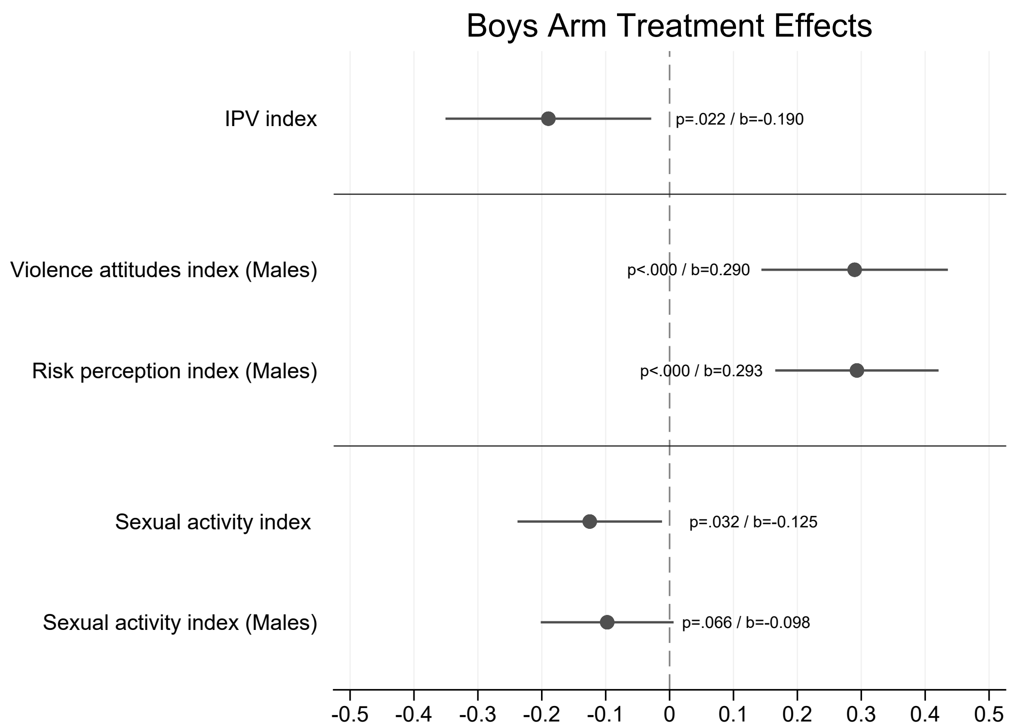


Figure 4 ITT Effects of *Boys* Treatment

Notes. This figure presents estimates of β_1 from equation 1 for separate regressions with the outcome specified on the y-axis. Outcomes have been standardized so that the x-axis is in standard deviation units. Outcomes are centered around females (males) in ELA only communities who were not assigned (whose connected female was not assigned) to the *Goal* treatment. Bolded markers are statistically significant at $p < 0.1$. p-values and coefficient estimates are displayed beside each marker.
Source. Female and male baseline and endline data, balanced panel.

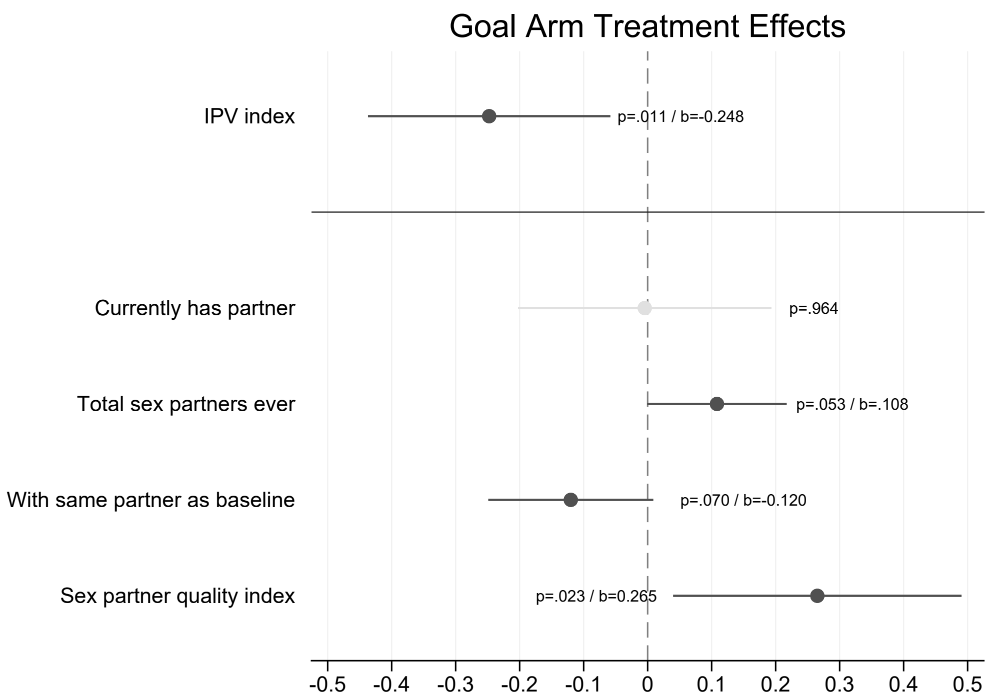


Figure 5 ITT Effects of *Goal* Treatment

Notes. This figure presents estimates of β_2 from equation 1 for separate regressions with the outcomes specified on the y-axis. Outcomes have been standardized so that the x-axis is in standard deviation units. Outcomes are centered around females (males) in ELA only communities who were not assigned (whose connected female was not assigned) to the *Goal* treatment. Bolded markers are statistically significant at $p < 0.1$. p-values and coefficient estimates are displayed beside each marker.
Source. Female and male baseline and endline data, balanced panel.

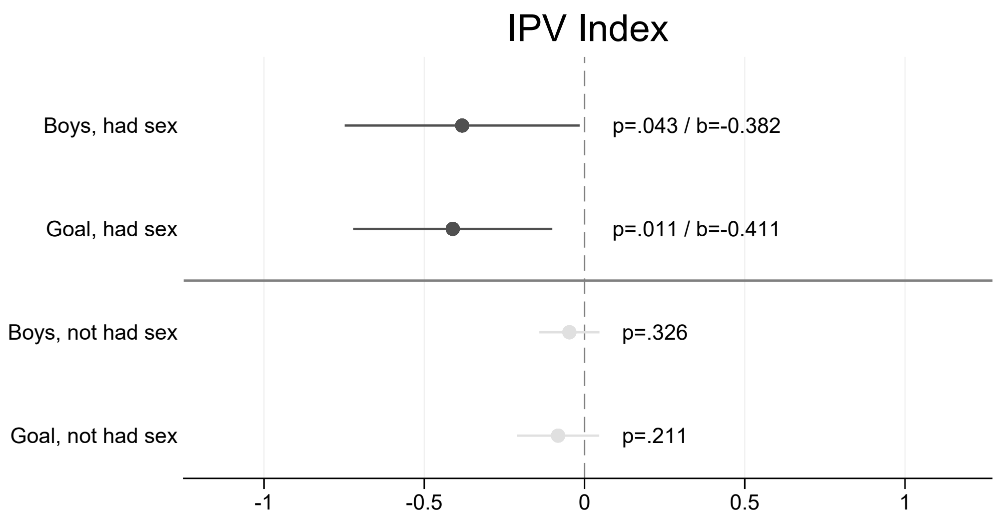


Figure 6 Impact of Treatments on IPV, Heterogeneity Sexually Active at Baseline

Notes. This figure presents estimates of β_1 and β_2 from equation 1 splitting the data by sexual activity status at baseline. The IPV index is centered on females in ELA only communities who were sexually active at baseline and who were not assigned to the *Goal* treatment. Bolded markers are statistically significant at $p < 0.1$. p-values and coefficient estimates are displayed beside each marker.
Source. Female baseline and endline data, balanced panel. The sample in the top half is restricted to females who are sexually active at baseline. The sample in the bottom half is restricted to females who are not sexually active at baseline.

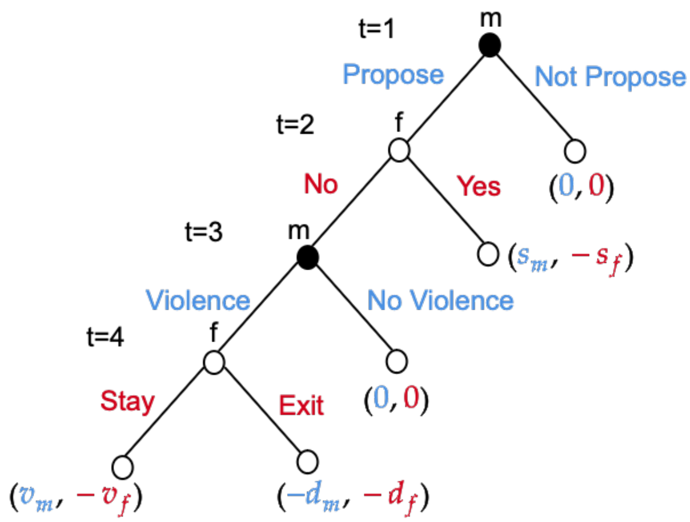


Figure 7 Game Tree

Notes. This figure presents the complete information version of the game to give the reader a flavor of the game. This is a one-shot, sequential game, indicated by $t = 1, 2, 3, 4$, where players know their own payoffs but do not know each other's payoffs. Blue text denotes male strategies and payoffs and red text denotes female strategies and payoffs, and m indicates a male decision node and f indicates a female decision node.

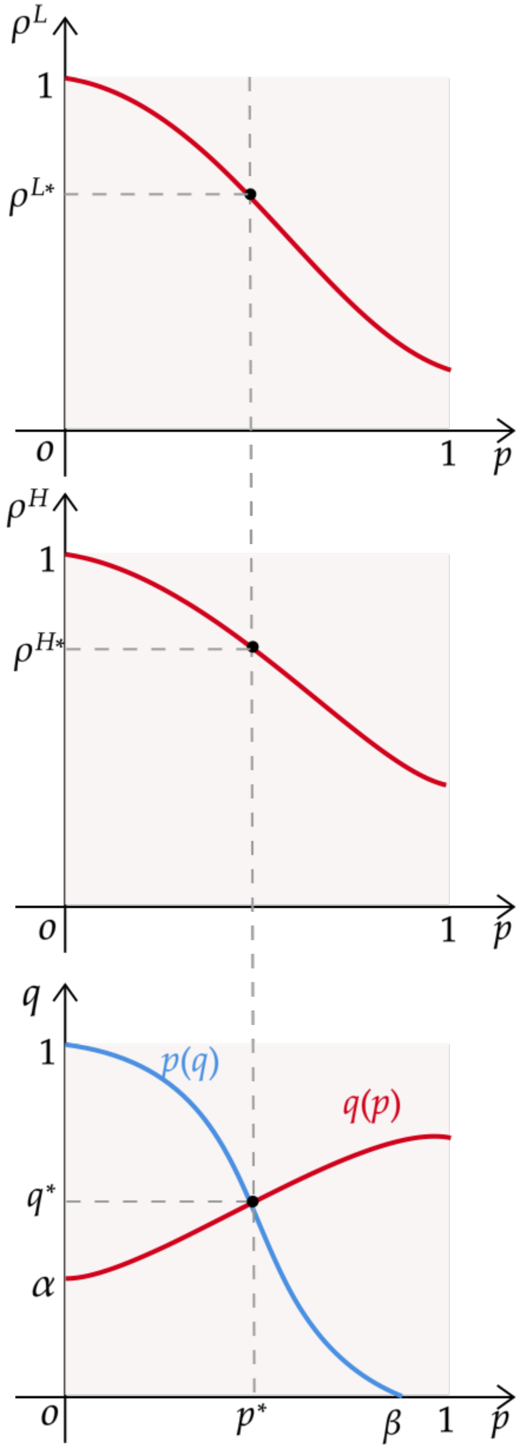


Figure 8 Equilibrium

Notes. This figure presents the model equilibrium as described in Proposition 1. The top panel presents the best response function for low-type females, the middle panel presents the best response function for high-type females, and the bottom panel shows the unique equilibrium point given by the crossing point of the $q(p)$ curve and the best response function for males, the $p(q)$ curve.

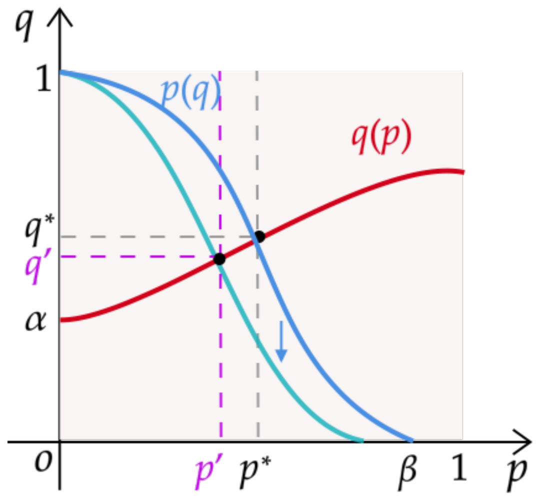
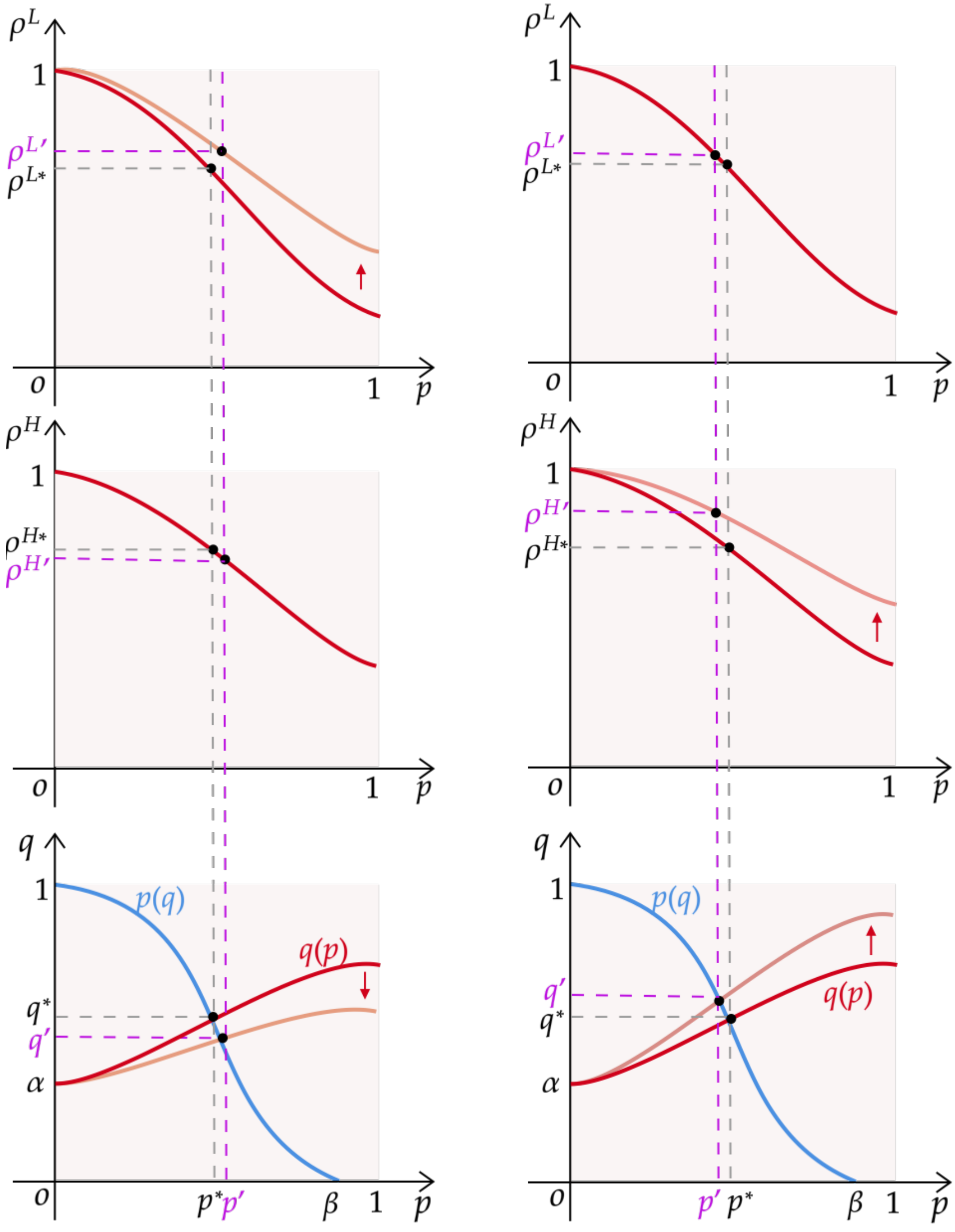


Figure 9 A Change in Males' Payoff from Violence

Notes. This figure demonstrates the dynamic impacts of a decrease in the net benefit of violence for males, v_m, on the model equilibrium. The males' best response function, $p(q)$ will shift inward, resulting in a lower equilibrium p' and q'.



(a) Low-type females

(b) High-type females

Figure 10 A Shift in Females' Payoff from Risky Sex

Notes. This figure demonstrates the dynamic impacts of an increase in the cost of risky sex for females, s_f on the model equilibrium. Panel (a) demonstrates the impacts if only low-types' s_f increases and Panel (b) demonstrates the impacts if only high-types' s_f increases. In Panel (a), low-type females become more likely to say no for every value of p, shown in the top panel. This causes $q(p)$ shift downward, shown in the bottom panel, resulting in a lower equilibrium q' and a higher equilibrium p'. In Panel (b), high-type females become more likely to say no for every value of p, shown in the middle panel. This causes $q(p)$ shift upward, shown in the bottom panel, resulting in a higher equilibrium q' and a lower equilibrium p'.

Table 1 Treatment-Control Balance at Baseline

	(1)	(2)	(3)	(4)
Outcome	**ELA Only Control Mean**	**Boys-ELA**	**No Goal Control Mean**	**Goal - No Goal**
A. Intimate Partner Violence				
Psychological abuse often	0.017	0.006	0.018	0.006
		(0.008)		(0.006)
Psychological abuse in last year	0.054	0.017	0.062	0.006
		(0.017)		(0.010)
Physical abuse often	0.008	0.010	0.011	0.000
		(0.007)		(0.004)
Physical abuse in last year	0.045	0.011	0.053	-0.007
		(0.016)		(0.009)
Forced sex often	0.012	0.004	0.013	-0.001
		(0.006)		(0.004)
Forced sex in last year	0.035	0.006	0.040	-0.007
		(0.013)		(0.008)
B. Sexual Activity				
Ever had sex	0.250	-0.001	0.261	0.006
		(0.035)		(0.018)
Currently has a partner	0.212	0.011	0.230	-0.006
		(0.034)		(0.017)
Had a partner in the past 2 years	0.266	0.010	0.279	0.003
		(0.037)		(0.019)
Total sex partners ever	0.312	0.018	0.334	0.001
		(0.054)		(0.026)
Hours with boyfriend in the past day	0.030	0.014	0.041	-0.002
		(0.013)		(0.011)
C. Demographic Characteristics				
Never talks to mother about SRH	0.839	-0.004	0.830	-0.001
		(0.022)		(0.016)
Age in years	16.45	-0.625*	16.18	0.045
		(0.336)		(0.115)
Highest grade attended	8.01	-0.182	8.04	-0.059
		(0.283)		(0.107)
Married or cohabiting	0.074	-0.010	0.078	0.000
		(0.019)		(0.010)
Household owns their house	0.674	-0.002	0.653	-0.014
		(0.039)		(0.018)
Number of household members	3.28	-0.023	3.27	-0.043
		(0.130)		(0.048)
Observations	1,074	3,178	2,313	3,178
χ^2 p-value		.535		.867

Notes. Column 1 shows means for females in ELA only communities and column 3 shows means for females not assigned to the *Goal* treatment. Columns 2 and 4 test for differences between the means in the community- or individual-level treatment arms and the corresponding control group means, controlling for the randomization strata. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates in columns 2 and 4. The χ^2 p-value in the last row is the p-value from a test of the joint significance of all outcomes in Panel A and B. ***p<.01, **p<.05, *p<.10.
Source. Female baseline data.

Table 2 Impact of Treatments on Intimate Partner Violence (IPV)

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
IPV index					
treatment x post	-0.190**	-0.248***	0.608	0.000	5,182
	(0.082)	(0.097)			
Psychological abuse often					
treatment x post	-0.011	-0.012	0.941	0.026	5,182
	(0.012)	(0.013)			
Psychological abuse in last year					
treatment x post	-0.029	-0.024	0.840	0.086	5,182
	(0.024)	(0.030)			
Physical abuse often					
treatment x post	-0.018**	-0.020*	0.827	0.019	5,182
	(0.008)	(0.011)			
Physical abuse in last year					
treatment x post	-0.037*	-0.019	0.496	0.062	5,182
	(0.019)	(0.026)			
Force sex often					
treatment x post	-0.028**	-0.035***	0.588	0.023	5,182
	(0.011)	(0.013)			
Force sex in last year					
treatment x post	-0.028*	-0.059**	0.206	0.045	5,182
	(0.017)	(0.024)			

Notes. This table presents estimates of β_1 and β_2 from equation 1. For each outcome, the coefficients from a single regression are presented in a row, with estimates of β_1 in column 1 and β_2 in column 2. Column 3 presents the p-value for a test of whether β_1 is equal to β_2. Column 4 presents the control mean at endline and column 5 shows the number of observations in the model. All specifications include controls for highest grade attended, whether the female's household owns the house she lives in, whether the female talks to her mom about sexual reproductive health topics, age of the female, and ELA club fixed effects. IPV index is generated by taking the unweighted mean across the six IPV indicators after they have been standardized to the mean and standard deviation among females in the control group at baseline and endline separately. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel.

Table 3 Impact of Treatments on Male IPV and SRH Attitudes (Male Data)

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
A. Male Violence Attitudes					
Violence attitudes index					
treatment x post	0.290***	0.066	0.024	0.000	2,314
	(0.074)	(0.099)			
Women should not tolerate violence from husband/partner					
treatment x post	0.174***	0.014	0.001	0.727†	2,314
	(0.035)	(0.048)			
Men should not beat women under any circumstances					
treatment x post	0.039	0.025	0.812	0.831	2,314
	(0.043)	(0.058)			
B. Male Risk Perception					
Risk perception index					
treatment x post	0.293***	-0.022	0.001	0.000	2,314
	(0.065)	(0.098)			
Male believes female friend is somewhat or very likely to have STI					
treatment x post	0.140***	0.053	0.184	0.581	2,314
	(0.048)	(0.066)			
Male believes that over 15% of males in his community have STIs					
treatment x post	0.123***	-0.019	0.035	0.308	2,314
	(0.046)	(0.066)			
Girls have right to ask to use condom					
treatment x post	0.122***	-0.066	0.001	0.797	2,314
	(0.043)	(0.059)			
C. Male Sexual Activity					
Sexual activity index					
treatment x post	-0.098*	-0.006	0.195	-0.062	2,314
	(0.053)	(0.072)			
Ever had sex					
treatment x post	-0.028	-0.006	0.587	0.423	2,314
	(0.030)	(0.041)			
Currently has a partner					
treatment x post	-0.060*	-0.006	0.299	0.345	2,314
	(0.036)	(0.053)			
Total sex partners ever					
treatment x post	-0.116**	0.003	0.045	0.398	2,314
	(0.047)	(0.060)			

Notes. This table presents estimates of β_1 and β_2 from equation 10. For each outcome, the coefficients from a single regression are presented in a row, with estimates of β_1 in column 1 and β_2 in column 2. Column 3 presents the p-value for a test of whether β_1 is equal to β_2. Column 4 presents the control mean at endline and column 5 shows the number of observations in the model. All specifications include controls for age of the male, highest grade attended, a binary indicator that the male never talks to his father about sexual reproductive health topics, a binary indicator that the male's household owns the house he lives in, and region fixed effects. The indexes in Panels A, B, and C are the unweighted mean of the indicators that follow after they have been standardized to the mean and standard deviation among males in the control group at baseline and endline separately. Standard errors, clustered at the attached female level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.

†Baseline mean in the *Boys* treatment arm. At baseline, 72.7% of males in the *Boys* treatment arm agreed with this statement compared to 85.0% in the males' control group. At endline, these means had changed to 93.2% and 88.6%, respectively, generating the treatment effect in the *Boys* arm.

Source. Male baseline and endline data, balanced panel.

Table 4 Impact of Treatment on Partner Churn and Quality of Sex Partners

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
A. Sexual Activity					
Sexual activity index					
treatment x post	-0.125**	0.065	0.004	0.000	5182
	(0.058)	(0.064)			
Ever had sex					
treatment x post	-0.047	0.046	0.010	0.372	5,182
	(0.031)	(0.031)			
Currently has a partner					
treatment x post	-0.110***	0.000	0.007	0.337	5,182
	(0.036)	(0.044)			
Total sex partners ever					
treatment x post	-0.025	0.088**	0.015	0.491	5,182
	(0.037)	(0.038)			
B. Partner Churn					
With same partner as baseline					
treatment x post	-0.020	-0.039*	0.468	0.117	2,591
	(0.021)	(0.021)			
C. Partner Quality					
Quality index					
treatment x post	-0.052	0.265**	0.170	0.000	1,711
	(0.106)	(0.114)			
His age					
treatment x post	0.114	-0.488	0.353	24.998	1,711
	(0.537)	(0.410)			
Dropout/never enroll					
treatment x post	-0.007	-0.047**	0.141	0.041	1,711
	(0.023)	(0.023)			

Notes. Panels A and C present estimates of β_1 and β_2 from equation 1. In Panel B, the reported coefficients are β_1 and β_2 from an adapted version of equation 1 that uses only one round of data, where the outcome is the change in partnership status from baseline to endline. In Panel C, His Age is the average age in years of all sexual partners listed and Dropout/Never Enrolled is the share of sexual partners listed whose enrollment status is dropped out or never enrolled in school.The Quality index is generated by taking the unweighted mean of the indicators after they have each been recoded so that positive coefficients indicate improved outcomes and standardized to the mean and standard deviation among females control group at baseline and endline separately. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10
Source. Panel A: Female baseline and endline data, balanced panel. Panel B: Female data, outcome from endline, controls from baseline, balanced panel. Panel C: Female baseline and endline data, balanced panel. Sample is restricted to females in the balanced panel who list at least one sexual partner at baseline or endline.

Table 5 Attrition: Female Sample

	(1)	(2)	(3)	(4)
	×Boys×Goal	**×Boys**	**×Goal**	**Levels**
Panel A. Differential Attrition by Treatment Status, fully-interacted				
Boys			-0.031	-0.012
			(0.038)	(0.032)
Goal				0.005
				(0.025)
Panel B. Differential Attrition by Treatment Status and Key Measures				
Never talks to mother about SRH	0.033	-0.014	-0.056	0.001
	(0.094)	(0.050)	(0.078)	(0.041)
Age in years	-0.010	0.008	-0.001	-0.000
	(0.012)	(0.010)	(0.008)	(0.007)
Highest grade attended	0.019	-0.016	-0.003	0.006
	(0.014)	(0.010)	(0.010)	(0.008)
Married or cohabiting	-0.076	-0.011	-0.082	0.021
	(0.101)	(0.085)	(0.065)	(0.065)
Household owns their house	0.091	-0.058	-0.076	0.027
	(0.080)	(0.041)	(0.054)	(0.029)
House has electricity	0.057	-0.027	0.011	0.020
	(0.066)	(0.045)	(0.040)	(0.030)
Number of household members	-0.044	-0.004	0.009	-0.008
	(0.028)	(0.016)	(0.019)	(0.011)
Boys			0.004	0.077
			(0.196)	(0.154)
Goal				0.108
				(0.112)
Observations		3,178		

Notes. Each panel presents coefficients from a single regression where the outcome, Y_{ic}, is an indicator equal to 1 if the female was not resurveyed at endline (i.e., attrited). In each panel, the rows list demographic variables included in the model, and the columns indicate interaction terms. Column 1 presents coefficient estimates on interactions between the row variables and a treatment indicator of being assigned to both the *Boys* and *Goal* treatment, column 2 presents coefficient estimates on interactions of the row variables with being assigned to the *Boys* treatment, column 3 presents coefficient estimates on interactions of the row variables with being assigned to the *Goal* treatment, and column 4 presents coefficient estimates on the row variables. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Female baseline data.

Table 6 Attrition: Male Sample (Male Data)

	(1)	(2)	(3)	(4)
	×Boys×Goal	**×Boys**	**×Goal**	**Levels**
Panel A. Differential Attrition by Treatment Status, fully-interacted				
Boys			-0.001	-0.006
			(0.033)	(0.022)
Goal				-0.009
				(0.027)
Panel B. Differential Attrition by Treatment Status and Key Measures				
Never talk to dad about SRH	-0.075	0.028	0.099	-0.070
	(0.075)	(0.063)	(0.065)	(0.057)
Age	0.011	0.006	-0.009	-0.005
	(0.009)	(0.005)	(0.008)	(0.005)
Highest grade attended	0.011	-0.018*	-0.002	0.011
	(0.012)	(0.009)	(0.010)	(0.008)
Married or cohabiting	-0.062	-0.096	0.142	0.048
	(0.160)	(0.060)	(0.146)	(0.046)
Household owns their house	0.030	-0.066	-0.028	0.038
	(0.082)	(0.050)	(0.067)	(0.044)
House has electricity	-0.154**	0.040	0.087	-0.035
	(0.078)	(0.043)	(0.056)	(0.036)
Number of household members	0.032	0.006	-0.010	-0.000
	(0.027)	(0.016)	(0.022)	(0.013)
Boys			-0.227	0.025
			(0.190)	(0.112)
Goal				0.063
				(0.158)
Observations		1,466		

Notes. Each panel presents coefficients from a single regression where the outcome, Y_{ic}, is an indicator equal to 1 if the male was not resurveyed at endline. In each panel, the rows list demographic variables included in the model, and the columns indicate interaction terms. Column 1 presents coefficient estimates on interactions between the row variables and a treatment indicator of being assigned to both the *Boys* and *Goal* treatment, column 2 presents coefficient estimates on interactions of the row variables with being assigned to the *Boys* treatment, column 3 presents coefficient estimates on interactions of the row variables with being assigned to the *Goal* treatment, and column 4 presents coefficient estimates on the row variables. Standard errors, clustered at the attached female level, are presented in parentheses below the coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Male baseline data.

Appendices

A Figures and Tables

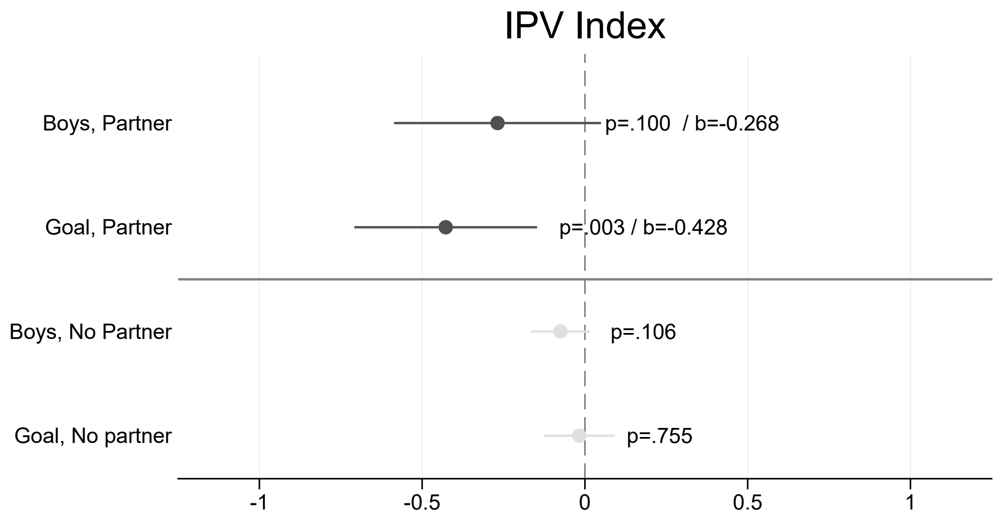


Figure A1 Impact of Treatments on IPV, Heterogeneity by Had Partner in Past Two Years at Baseline

Notes. This figure presents estimates of β_1 and β_2 from equation 1. The IPV index is centered on females in the ELA only communities who had a partner in the past two years at baseline and were not assigned to the *Goal* treatment. Bolded markers are statistically significant at $p < 0.1$. p-values and coefficient estimates are displayed beside each marker.
Source. Female baseline and endline data, balanced panel. The sample in the top half is restricted to females who had a partner in the past two years at baseline. The sample in the bottom half is restricted to females who did not have a partner in the past two years at baseline.

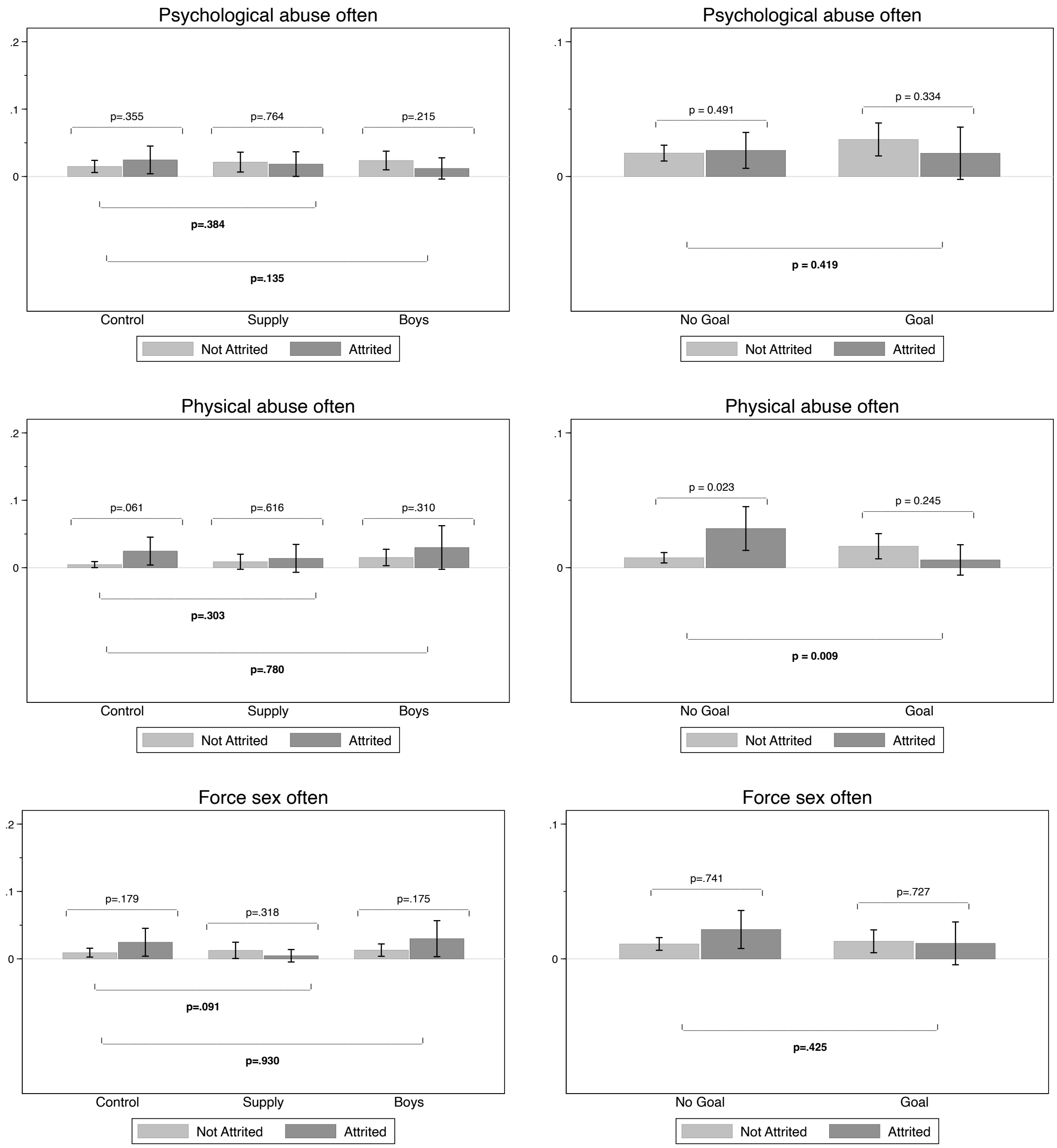


Figure A2 Baseline Outcomes by Attrition Status: IPV

Notes. This figure shows the mean of each outcome among females who attrited and those who did not by treatment status. The left-side column presents these means for each community-level treatment arm, specified on the x-axis. Above the mean bars are p-values for tests of equality between the means of attrited and non-attrited females. Below the mean bars are p-values for tests that attrition differed between treatment arms, first comparing the *Supply* arm to ELA only communities and then comparing the *Boys* arm to the ELA only communities. The right-side column presents the outcome means by attrition status for females assigned to the *Goal* treatment and those who were not, separately. Above the mean bars are p-values for tests of equality of means between attrited and non-attrited femaleswithin each arm. Below the mean bars is a p-value from a test that attrition differed between females who were assigned to *Goal* and those who were not.
Source. Female baseline data.

Table A1 Intimate Partner Violence and Risky Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	Female Data			**Male Data**		
	Use condom with last partner	**Use condom always**	**Number of sex partners**	**Use condom with last partner**	**Use condom always**	**Number of sex partners**
IPV often	-0.097*	-0.040	-0.020	0.014	-0.108*	0.447**
	(0.059)	(0.048)	(0.081)	(0.077)	(0.057)	(0.193)
IPV in last year	0.019	-0.028	0.076	-0.091*	-0.147***	0.393**
	(0.040)	(0.031)	(0.070)	(0.051)	(0.035)	(0.156)
Psychological abuse often	-0.057	0.009	-0.164	-0.070	-0.184***	0.608**
	(0.071)	(0.056)	(0.109)	(0.090)	(0.049)	(0.254)
Physical abuse often	-0.221**	-0.044	-0.044	0.186	-0.041	0.201
	(0.090)	(0.066)	(0.117)	(0.181)	(0.180)	(0.184)
Force Sex often	-0.059	-0.012	0.160	0.101	0.119	0.073
	(0.090)	(0.080)	(0.129)	(0.123)	(0.130)	(0.130)
Psychological abuse in last year	0.001	-0.018	0.001	-0.108*	-0.178***	0.497**
	(0.043)	(0.038)	(0.094)	(0.057)	(0.035)	(0.213)
Physical abuse in last year	-0.008	-0.046	0.103	-0.124**	-0.105**	0.526**
	(0.047)	(0.039)	(0.102)	(0.058)	(0.043)	(0.241)
Force sex in last year	-0.040	-0.064	0.126	-0.065	-0.116***	0.379*
	(0.054)	(0.040)	(0.127)	(0.058)	(0.042)	(0.225)
Observations	800	800	800	678	678	678

Notes. Each cell of this table presents a coefficient from a separate regression of intimate partner violence (defined in the rows) on condom use (columns 1 and 4, with last partner and columns 2 and 5, always use) and the number of sexual partners in the last six months (columns 3 and 6). Standard errors are clustered at the club level in columns 1–3 and at the attached female level in columns 4–6 and are presented in parentheses below the coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Columns 1–3: Female baseline data, restricted to females who were sexually active at baseline. Columns 4–6: Male baseline data, restricted to males who were sexually active at baseline.

Table A2 Predictors of Boys Participation in Boys Soccer Treatment

	(1)	(2)
	Contacted by GRS	**Enrolled in GRS**
Enrolled in school	0.076**	0.133***
	(0.037)	(0.051)
Household owns their house	0.037	-0.054
	(0.037)	(0.037)
Never talk to dad about SRH	-0.026	-0.000
	(0.049)	(0.050)
Age in years	0.027***	0.001
	(0.004)	(0.005)
House has electricity	-0.021	0.033
	(0.040)	(0.042)
Number of household members	0.001	0.040***
	(0.015)	(0.015)
Iringa	-0.246***	0.078
	(0.045)	(0.050)
Mbeya	-0.084*	0.131***
	(0.048)	(0.050)
Outcome mean	0.679	0.352
Observations	787	787

Notes. This table shows the results from two regressions, present in each column. In column 1, the outcome is an indicator that the GRS team contacted the male in our study sample; in column 2, the outcome is an indicator that the GRS team enrolled the male in our study sample. Covariates are shown in the rows. Standard errors, clustered at the attached female level, are presented in parentheses below the coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Male baseline data, restricted to males in the *Boys* treatment arm.

Table A3 Predictors of Achieving Goal

	(1)	(2)
	Number of Strategies Set	**Number of Strategies Achieved**
Major depressive disorder likely	-0.235*	-0.219*
	(0.121)	(0.127)
Self-efficacy	0.103***	0.102***
	(0.036)	(0.035)
Will take risk (top quintile)	0.055	-0.002
	(0.057)	(0.067)
Impatient now and patient later	-0.028	-0.018
	(0.076)	(0.082)
Age in years	-0.002	0.008
	(0.009)	(0.013)
Currently enrolled in school	-0.185	-0.168
	(0.114)	(0.135)
Completed school	-0.093	-0.054
	(0.106)	(0.125)
Household has electricity	-0.044	0.021
	(0.051)	(0.055)
Household has earthen floor	-0.232***	-0.121
	(0.065)	(0.088)
Outcome mean	1.772	1.640
Observations	789	644

Notes. The outcome in column 1 is the number of strategies the female set and committed to during the goal setting activity, ranging from 1 to 3 strategies; the outcome in column 2 is the number of strategies that the female reports she achieved at the endline survey, ranging from 0 to 3 strategies. The covariates are from the baseline survey (age, currently enrolled in school, completed school, household has electricity, household has earthen floor) or the goal setting activity, 1st visit (major depressive disorder likely, self-efficacy, will take risk). Major depressive disorder likely is a binary indicator measured using the Patient Health Questionnaire-2 (PHQ-2), where a score of 3 or higher is indicative of depression. Self-efficacy is measured using the General Self-Efficacy Scale developed by Schwarzer and Jerusalem (1995). This scale generates a total self-efficacy score that ranges from 10-40, which we standardized using the mean and standard deviation among females in ELA only communities. Will take risk is based on a question that asked females how willing they were to take risks on a scale from 1 to 10. We generate a binary indicator equal to one if the female gave an answer in the top quintile of responses. The excluded category for schooling is dropout or never enrolled. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Column 1: Female goal setting participants, 1st visit. Column 2: Female goal setting participants, endline data.

Table A4 External Validity: Comparison of Our Sample to a Random Sample

	(1)	(2)	(3)	(4)	(5)
	Female Baseline Mean	**Buehren et al. (2017) Baseline Mean**	**Buehren et al. (2017) Baseline Control Mean**	**(1)-(2) p-value**	**(1)-(3) p-value**
Age in year	16.355	16.676	16.621	.023	.097
Married	0.044	0.022	0.024	.024	.052
Household owns their house	0.605	0.606	0.626	.951	.560
Has child	0.089	0.068	0.073	.115	.283
Ever had sex	0.241	0.228	0.242	.557	.948
Uses condom	0.182	0.188	0.189	.739	.765
Highest grade attended	8.273	8.342	8.239	.648	.835
Dropout from school	0.236	0.212	0.217	.243	.438
Never enrolled in school	0.012	0.015	0.019	.518	.234
Observations	1,621	4,954	1,708		

Notes. Column 1 shows baseline means among females in Dodoma and Iringa in the current study sample, collected in 2016. Columns 2 and 3 show baseline means (before ELA clubs were established) among females in the Buehren et al. (2017) study, collected in 2009. The Buehren et al. (2017) sample comprises a random sample of adolescent females from the same communities as the current study in Dodoma and Iringa. Buehren et al. (2017) treatment communities are where ELA clubs were eventually established. Column 2 shows means using the entire baseline sample from Buehren et al. (2017). Column 3 shows means using baseline data restricted to the Buehren et al. (2017) control communities. Column 4 presents the p-values from a test of equality between the means in columns 1 and 2. Column 5 presents the p-values from a test of equality between the means in columns 1 and 3. The latter comparison is made to alleviate concern that communities where clubs are established are different from surrounding communities.
Source. Column 1: Female baseline data, restricted to females in Dodoma and Iringa. Columns 2–3: Buehren et al. (2017) female baseline data. In columns 1–3, the sample is restricted to females aged 13-22 in order for their to be a common age range across samples.

Table A5 External Validity: IPV Reports, Females Aged 15-24

	(1)	(2)
	Female Baseline Mean	**DHS 2016 Mean**
Panel A. Partnered and unpartnered females		
IPV often	0.040	0.049
IPV in last year	0.115	0.154
Psychological abuse often	0.027	0.037
Physical abuse often	0.013	0.012
Force sex often	0.015	0.016
Psychological abuse in last year	0.088	0.135
Physical abuse in last year	0.069	0.049
Force sex in last year	0.051	0.053
Jealous or angry if talk to other men	0.313	0.281
Insists on knowing where you are	0.277	0.271
Sample size	2,140	3,275
Panel B. Partnered females		
IPV often	0.097	0.104
IPV in last year	0.282	0.322
Psychological abuse often	0.066	0.077
Physical abuse often	0.031	0.026
Force sex often	0.038	0.034
Psychological abuse in last year	0.216	0.283
Physical abuse in last year	0.168	0.103
Force sex in last year	0.124	0.111
Jealous or angry if talk to other men	0.764	0.600
Insists on knowing where you are	0.677	0.568
Sample size	876	1,820

Notes. This table presents average reports of IPV by females ages 15-25 at baseline in column 1, by females ages 15-24 as reported in the 2015 Tanzania DHS data in column 2. IPV items from the DHS were selected to match the specific items used in the study surveys. Psychological violence is measured as being threatened with harm and being insulted or made to feel bad. Physical violence is measured as being pushed, shaken, or throwing something. Force sex is measured as being forced to engage in sexual intercourse. Each item is reported as occurring often, sometimes, or not in the last 12 months. For each violence type, we generate indicators for the violence being reported as occurring often or at all in the last 12 months (in last year). IPV often is an indicator equal to one if the female report experiencing any of the IPV items (psychological, physical, and sexual) occurring often. IPV in last year is an indicator equal to one if the female reports experiencing any of the IPV items (psychological, physical, and sexual) occurring in the last 12 months.
Source. Column (1) Female baseline data for females ages 15-24, balanced panel. (2) Tanzania DHS 2016 female data for females ages 15-24.

Table A6 Impact of Treatments on Intimate Partner Violence (IPV), No Controls

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
IPV index					
treatment x post	-0.181**	-0.264***	0.457	0.000	5,182
	(0.082)	(0.097)			
Psychological abuse often					
treatment x post	-0.011	-0.014	0.813	0.026	5,182
	(0.012)	(0.013)			
Psychological abuse in last year					
treatment x post	-0.025	-0.030	0.867	0.086	5,182
	(0.024)	(0.030)			
Physical abuse often					
treatment x post	-0.017**	-0.021*	0.697	0.019	5,182
	(0.008)	(0.011)			
Physical abuse in last year					
treatment x post	-0.034*	-0.024	0.694	0.062	5,182
	(0.020)	(0.026)			
Force sex often					
treatment x post	-0.027**	-0.036***	0.515	0.023	5,182
	(0.011)	(0.013)			
Force sex in last year					
treatment x post	-0.026	-0.063***	0.144	0.045	5,182
	(0.017)	(0.024)			

Notes. This table presents estimates of β_1 and β_2 from equation 1. For each outcome, the coefficients from a single regression are presented in a single row, with estimates of β_1 in column 1 and β_2 in column 2. Column 3 presents the p-value for a test of whether β_1 is equal to β_2. Column 4 presents the control mean at endline, and column 5 shows the number of observations in the model. All specifications include club fixed effects. IPV index is generated by taking the unweighted mean across the six indicators after they have been standardized to the mean and standard deviation among females in the control group at baseline and endline separately. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel.

Table A7 Impact of Treatment on Intimate Partner Violence (IPV), Separating Goal into High and Low Strategies

	(1)	(2)	(3)	(4)	(5)	(6)
	Boys Treatment	**2-3 strategy Treatment**	**1 strategy Treatment**	**2-3 strategy = 1 strategy p-value**	**Endline Control Mean**	**Observations**
IPV index						
treatment x post	-0.190**	-0.323***	-0.131	0.065	0.000	5,182
	(0.082)	(0.111)	(0.102)			

Notes. This table presents estimates of a modified specification of equation 1 that splits the goal treatment indicator into two mutually exclusive categories for females assigned to the *Goal* arm: set 2 or 3 strategies or set 0 or 1 strategy. For each outcome, the coefficients from a single regression are presented in a single row, with estimates of β_1 in column 1 and the coefficients on the two *Goal* arm indicators in columns 2 and 3. Column 4 presents the p-value for a test of whether column 1 is equal to column 2. Column 5 presents the control mean at endline, and column 6 shows the number of observations in the model. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel.

Table A8 Impact of Treatments on Intimate Partner Violence (IPV), Goal Interactions

	(1)	(2)	(3)	(4)	
	Boys Treatment	**Goal Treatment**	**Boys = Goal treatment p-value**	**Endline Control Mean**	**Observations**
IPV index					
treatment x post	-0.190**	-0.248***	0.608	0.000	5,182
	(0.082)	(0.097)			
treatment x post x Goal	0.110				
	(0.173)				
Psychological abuse often					
treatment x post	-0.011	-0.012	0.941	0.026	5,182
	(0.012)	(0.013)			
treatment x post x Goal	-0.024				
	(0.020)				
Psychological abuse in last year					
treatment x post	-0.029	-0.024	0.840	0.086	5,182
	(0.024)	(0.030)			
treatment x post x Goal	-0.020				
	(0.042)				
Physical abuse often					
treatment x post	-0.018**	-0.020*	0.827	0.019	5,182
	(0.008)	(0.011)			
treatment x post x Goal	0.003				
	(0.018)				
Physical abuse in last year					
treatment x post	-0.037*	-0.019	0.496	0.062	5,182
	(0.019)	(0.026)			
treatment x post x Goal	0.007				
	(0.034)				
Force sex often					
treatment x post	-0.028**	-0.035***	0.588	0.023	5,182
	(0.011)	(0.013)			
treatment x post x Goal	0.035**				
	(0.014)				
Force sex in last year					
treatment x post	-0.028*	-0.059**	0.206	0.045	5,182
	(0.017)	(0.024)			
treatment x post x Goal	0.057**				
	(0.028)				

Notes. This table presents estimates of β_1, β_2, and γ_1 from equation 1. For each outcome, the coefficients from a single regression are presented in two sub-rows, with estimates of β_1 in column 1 and β_2 in column 2, labeled treatment x post, and the estimate of γ_1 in the second of the two rows, labeled treatment x post x Goal. Column 3 presents the p-value for a test of whether β_1 is equal to β_2. Column 4 presents the control mean at endline, and column 5 shows the number of observations in the model. All specifications include controls for highest grade attended, whether the female's household owns the house she lives in, whether the female talks to her mom about sexual reproductive health topics, age of the female, and club fixed effects. IPV index is generated by taking the unweighted mean across the six indicators after they have been standardized to the mean and standard deviation among females in the control group at baseline and endline separately. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel.

Table A9 Impact of Treatments on Males, Restricted to Sexually Active at Baseline (Male Data)

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
A. Male Violence Attitudes					
Violence attitudes index					
treatment x post	0.413***	0.100	0.046	0.000	1,100
	(0.112)	(0.153)			
Women should not tolerate violence from husband/partner					
treatment x post	0.204***	0.057	0.050	0.741†	1,100
	(0.050)	(0.074)			
Men should not beat women under any circumstance					
treatment x post	0.069	0.008	0.415	0.786	1,100
	(0.062)	(0.077)			
B. Male SRH Attitudes					
Risk perception index					
treatment x post	0.239**	0.044	0.160	0.000	1,100
	(0.096)	(0.136)			
Female friend somewhat or very likely to have STI					
treatment x post	0.106	0.071	0.683	0.619	1,100
	(0.065)	(0.083)			
Over 15% of males have an STI					
treatment x post	0.056	-0.040	0.323	0.286	1,100
	(0.071)	(0.091)			
Girls have a right to ask to use a condom					
treatment x post	0.106*	-0.002	0.185	0.881	1,100
	(0.057)	(0.083)			

Notes. This table presents estimates of β_1 and β_2 from equation 10. See notes from Table 3 for detail on table structure, outcome definitions, and control variables. Standard errors, clustered at the attached female level, are presented in parentheses below coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.

†Baseline mean in the *Boys* treatment arm. At baseline, 74.1% of sexually active males in the *Boys* treatment arm agreed with this statement compared to 87.5% in the males' control group. At endline, these means had changed to 92.5% and 86.0%, respectively, generating the treatment effect in the *Boys* arm.

Source. Male baseline and endline data, balanced panel, but restricted to males who had ever had sex at baseline.

Table A10 Impact of Boys Treatment Among Enrolled

	(1)	(2)	(3)
	Boys Enrolled	**Endline Control Mean**	**Observations**
A. Female Data			
IPV index			
treatment x post	-0.300**	0.000	3,662
	(0.150)		
Sexual activity index			
treatment x post	-0.297***	0.000	3,662
	(0.060)		
B. Male Data			
Violence attitudes index			
treatment x post	0.442***	0.000	1,542
	(0.104)		
Risk perception index			
treatment x post	0.370***	0.000	1,542
	(0.087)		
Sexual activity index			
treatment x post	-0.097	0.000	1,542
	(0.073)		

Notes. The table presents estimates of β_1 in column 1 from estimating equation 1 in Panel A and equation 10 in Panel B after restricting the sample in the *Boys* treatment arm to females with males in their network who enrolled in GRS in Panel A and to males who enrolled in GRS in Panel B. Column 2 presents the control mean at endline, and column 3 shows the number of observations. See Tables 2 and 3 for outcome definitions and control variables. Standard errors are clustered at the club level in panel A and at the attached female level in panel B, and are presented in parentheses below coefficients estimates in column 1. ***p<.01, **p<.05, *p<.10.
Source. Panel A: Female baseline and endline data, balanced panel. Panel B: Male baseline and endline data, balanced panel. The sample in the *Boys* treatment arm is restricted to females with males in their network who enrolled in *Boys* treatment in Panel A and to males who enrolled in *Boys* treatment in Panel B.

Table A11 Impact of Treatments on Partnership Churn, By Baseline IPV

	(1)	(2)	(3)	(4)	(5)
	Boys Treatment	**Goal Treatment**	**Boys = Goal p-value**	**Endline Control Mean**	**Observations**
A. Experienced IPV at baseline					
With same partner as baseline					
treatment x post	-0.068	-0.305***	.003	0.395	204
	(0.019)	(0.021)			
B. Had not experienced IPV at baseline					
With same partner as baseline					
treatment x post	-0.021	-0.014	.788	0.097	2387
	(0.098)	(0.088)			
p-value	0.627	0.031			

Notes. This table presents estimates of β_1 and β_2 from equation 1. For each outcome, the coefficients from a single regression are presented in a single row, with estimates of β_1 in column 1 and β_2 in column 2. Column 3 presents the p-value for a test of whether β_1 is equal to β_2. Column 4 presents the control mean at endline, and column 5 shows the number of observations in the model. All specifications include controls for highest grade attended, whether the female's household owns the house she lives in, whether the female talks to her mom about sexual reproductive health topics, age of the female, and club fixed effects. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Panel A: Female baseline and endline data, balanced panel, restricted to females who experienced any IPV (psychological, physical, or sexual) at baseline. Panel B: Female baseline and endline data, balanced panel, restricted to females who did not experience any IPV at baseline.

Table A12 Endline Characteristics of Females, By Status of Baseline Partnership

	(1)	(2)	(3)
	Not with Baseline Partner	With Baseline Partner	*p*-value
A. Demographics			
Age in years	20.5	21.4	<.001
Highest grade attended	10.2	9.96	.220
Currently enrolled in school	0.223	0.124	.001
Married or cohabiting	0.267	0.396	.001
Household owns their house	0.572	0.520	.189
Number of household members	2.73	2.76	.753
B. Intimate Partner Violence			
IPV often	0.044	0.072	.147
IPV in last year	0.181	0.264	.014
Psychological abuse often	0.028	0.048	.201
Physical abuse often	0.019	0.032	.300
Force sex often	0.016	0.028	.333
Psychological abuse in last year	0.144	0.204	.051
Physical abuse in last year	0.095	0.160	.017
Force sex in last year	0.067	0.068	.978
Observations	430	250	

Notes. This table presents endline means of demographics (Panel A) and outcomes (Panel B) for females who are no longer with the same partner as at baseline in column 1 and who are with the same partner as at baseline in column 2. Column 3 presents the p-value on a test of differences between the means in columns 1 and 2. Robust standard errors are used in the calculation of p-values.
Source. Female endline data, balanced panel, restricted to females who name sexual partners at baseline.

Table A13 Secular Trends in IPV in ELA Only Communities

	(1)	(2)
	IPV index	
	Non-Goal	**Goal**
Post	-0.052	-0.323***
	(0.05)	(0.095)
Observations	1,285	458

Notes. Each column in this table presents coefficient estimates of θ_1 from a separate estimation of the following specification: $Y_{ic} = \alpha + \theta_1 Post_t + \psi age_i + \epsilon_{ic}$. In column 1, the sample is restricted to non-goal setting participants in ELA only communities. In column 2, the sample is restricted to goal setting participants in ELA only communities. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel, restricted to females in ELA only communities.

Table A14 Cost-Effectiveness Comparison

	(1)	(2)	(3)	(4)
	Cost per person	**Treatment effect**	**Type of violence**	**Cost per 0.25 SD effect**
Boys	$41	-0.190 SD	psychological, physical, sexual	$54
Goal	$38	-0.248 SD	psychological, physical, sexual	$38
Haushofer et al. (2019)				
Transfer to wives	$496	-0.26 SD	physical	$477
		-0.22 SD	sexual	$539
Transfer to husbands	$496	-0.18 SD	physical	$689
Hidrobo, Peterman and Heise (2016)				
Transfers to adult women	$240	-0.06 pp	physical/sexual	$400

Notes. Standard deviation is denoted SD and percentage point is denoted pp. Column 1 shows the per person cost of each treatment and column 2 reports the treatment effect as reported in the original paper. For Hidrobo, Peterman and Heise (2016) the 6 percentage point reduction was converted to a 0.25 standard deviation change by the authors.

B Balance Tables

Table B1 Impacts of *Supply* Treatment on Contraceptive Uptake

		(1)	(2)	(3)
Outcome	**Time**	**ELA Only Control Mean**	**Supply**	**DD**
Panel A. Whole Sample				
Injectable	Baseline	0.021	0.026	
	Endline	0.024	0.021	-0.008
				(0.008)
Implant	Baseline	0.007	0.011	
	Endline	0.008	0.022	0.010
				(0.007)
IUD	Baseline	0.007	0.003	
	Endline	0.002	0.003	0.005
				(0.004)
Observations	Baseline	1,074	1,012	
	Endline	1,766	1,746	5,598
Panel B. Balanced Panel Sample				
Injectable	Baseline	0.020	0.026	
	Endline	0.028	0.020	-0.014
				(0.010)
Implant	Baseline	0.008	0.010	
	Endline	0.011	0.023	0.009
				(0.009)
IUD	Baseline	0.007	0.004	
	Endline	0.000	0.005	0.008
				(0.006)
Observations	Baseline	871	795	
	Endline	871	795	3,332

Notes. Column 1 shows control means for females in ELA only communities and column 2 shows means for females assigned to the *Supply* treatment. Column 3 shows the simple difference-in-differences estimate between the *Supply* treatment and the ELA Only communities over time. Outcomes are indicators that the female used an injectable, implant, or intrauterine device (IUD) as contraception with her last partner. These are the primary methods that were promoted by the implementation partners in the *Supply* treatment. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates in column 3. ***p<.01, **p<.05, *p<.10.
Source. Female baseline data.

Table B2 Treatment-Control Balance at Baseline, Balanced Panel

	(1)	(2)	(3)	(4)
Outcome	**ELA Only Control Mean**	**Boys-ELA**	**No Goal Control Mean**	**Goal - No Goal**
A. Intimate Partner Violence				
Psychological abuse often	0.015	0.009	0.017	0.008
		(0.008)		(0.006)
Psychological abuse in last year	0.053	0.014	0.057	0.013
		(0.016)		(0.012)
Physical abuse often	0.005	0.011*	0.007	0.006
		(0.007)		(0.005)
Physical abuse in last year	0.040	0.010	0.044	0.008
		(0.015)		(0.011)
Forced sex often	0.009	0.004	0.011	0.001
		(0.006)		(0.005)
Forced sex in last year	0.032	0.001	0.033	0.004
		(0.011)		(0.009)
B. Sexual Activity				
Had sex	0.247	-0.006	0.252	0.020
		(0.038)		(0.021)
Has partner	0.204	0.013	0.218	0.006
		(0.036)		(0.020)
Had a partner in the past 2 years	0.258	0.009	0.266	0.019
		(0.040)		(0.021)
Total sexual partners ever	0.295	0.029	0.316	0.020
		(0.056)		(0.029)
Hours with boyfriend in the past day	0.030	0.019	0.044	0.002
		(0.014)		(0.013)
Panel C. Demographics				
Never talks to mother about SRH	0.844	-0.011	0.824	0.008
		(0.023)		(0.018)
Age in years	16.44	-0.715**	16.09	0.116
		(0.358)		(0.135)
Highest grade attended	7.99	-0.194	8.02	-0.118
		(0.301)		(0.119)
Married or cohabiting	0.075	-0.009	0.076	0.004
		(0.022)		(0.012)
Household owns their house	0.665	0.015	0.651	-0.016
		(0.040)		(0.021)
Number of household members	3.24	0.009	3.25	-0.055
		(0.135)		(0.055)
Observations	871	2,591	2,591	

Notes. Column 1 shows means for females in ELA only communities and column 3 shows means for females not assigned to the *Goal* treatment. Columns 2 and column 4 test for difference between means in the community- or individual-level treatment arms and the corresponding control group means, controlling for randomization strata. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates in columns 2 and 4. ***p<.01, **p<.05, *p<.10

Source. Female baseline data, balanced panel.

Table B3 Treatment-Control Balance at Baseline: Intimate Partner Violence, Balanced Panel, Sexually Active at Baseline

	(1)	(2)	(3)	(4)
Outcome	**ELA Only Control Mean**	**Boys-ELA**	**No Goal Control Mean**	**Goal - No Goal**
Psychological abuse often	0.060	0.032	0.059	0.039
		(0.033)		(0.029)
Psychological abuse in last year	0.169	0.095*	0.196	0.040
		(0.052)		(0.046)
Physical abuse often	0.010	0.054**	0.026	0.023
		(0.026)		(0.020)
Physical abuse in last year	0.109	0.089*	0.148	0.012
		(0.049)		(0.041)
Force sex often	0.020	0.032	0.037	-0.005
		(0.020)		(0.018)
Force sex in last year	0.085	0.037	0.102	0.000
		(0.041)		(0.029)
Observations	201	637	459	637

Notes. Column 1 shows means for females in ELA only communities and column 3 shows control for females not assigned to the *Goal* treatment. Columns 2 and column 4 test for difference between means in the community- or individual-level treatment arms and the corresponding control group means, controlling for randomization strata. Standard errors, clustered at the club level, are presented in parentheses below coefficient estimates in columns 2 and 4. ***p<.01, **p<.05, *p<.10
Source. Female baseline data, balanced panel, but restricted to females who had ever had sex at baseline.

C Grassroot Soccer Curriculum

The GRS curriculum comprised 11 one hour practices, including graduation. Ten of the 11 sessions were on SRH and one was on malaria. Each practice starts with a short "warm-up" where players recap the key messages of the previous practice and start discussing some of the themes of the current practice. The main part of the practice is a sport-based activity aimed at teaching and sparking discussions about healthy behaviors. Key messages are included throughout the activity. The practice ends with a quick "cool down" to recap the goals and key messages of the practice. In selected practices, the coaches also share personal stories about their real-life experiences related to the practice theme. Coaches are available post-practice for an additional 15-30 minutes in case males want one-on-one meetings to discuss more private issues.

Table C1 presents the goals and key messages of each practice.

Table C1 Grassroot Soccer Curriculum

(1)	(2)	(3)	(4)
Session	**Title**	**Goals**	**Sample key messages**
1	Join the Team!	Instill a core set of life skills and knowledge to help the boys stay healthy	"Build your team with strong supporters that help you abstain from sex or to practice safe sexual behaviors like using condoms."
2	Communicate	Develop good communication skills, in particular with girls	"Boys and girls can listen to each other and respect each other, even if it can be difficult." "When communicating with someone of the opposite sex, remember to: find a safe place to talk; show respect to the person you are communicating with; make strong eye contact; and stay positive." "In life, we should all stand up for girls and women to protect them from abuse."
3	Risky Partners	Emphasizes the skewed power dynamics and risks involved when a girl is in a sexual relationship with an older boy; Reinforces that abstinence is safest way to avoid HIV	"If you chose to have sex, it is much less dangerous to have sex with someone your own age than to have sex with someone older." "If you have less power in a relationship, it is harder to make healthy decisions, like using condoms or being mutually faithful."
4	Stop the Spread	Importance of HIV testing and safe sexual practices	"HIV spreads quickly when people have unprotected sex with multiple sexual partners."
5	Build your Team	Building supportive social networks	"We all need to build our team with strong supporters to stay strong in life."
6	One or None	Emphasizes importance of abstinence or having a single sexual partner to protect them from contracting HIV	"If you do choose to have sex, you can protect yourself by using condoms and having 1 mutually faithful partner that is HIV-negative."
7	Know the Game	Understanding how HIV attacks the body, living with HIV, and antiretroviral drugs	"The immune system protects the body from germs and diseases." "A healthy lifestyle can help an HIV-positive person live a longer, healthier life."
8	Protect yourself	Learn how to identify and use contraceptives, how condoms and circumcision protect from HIV	"You can say NO to unprotected sex!" "There are several ways to prevent pregnancy. Choose the one that's right for you and your partner"
9	Kick out Malaria	Discusses malaria, the use of bednets, and malaria treatment	"Malaria can kill. You can protect yourself from malaria by sleeping under a CCD every night."
10	Red Card	Uses the concept of "Red Card" in soccer to identify high risk situations to avoid, such as intimate partner violence, unprotected sex, old partners, multiple partners, and alcohol abuse	"Use the knowledge and skills you learned from the program to start conversations with friends and family members about difficult issues, like unprotected sex, older partners, multiple sexual partners, mixing sex and alcohol, and gender-based violence."
11	Make your Move!	Celebration of the boys' graduation based on the most important lessons	"Negative pressures from friends, family and the community can be difficult to avoid. But you can make your own smart choices in life!"

D Registered Outcomes

Empirical analysis is pre-registered through the AEA RCT registry (RCT ID: AEARCTR-0001305). In this appendix we present power calculations and estimation from our main specification on all registered outcomes.

For our study sample, we were limited to the total number of ELA clubs in Dodoma, Iringa, and Mbeya regions: 150 clubs with an average of 20 members. For our three arm study, this implies 50 clusters per treatment arm. We assume a 5% significance level, 80% power, an ICC of 0.1, and a coefficient of variation of 0.33. In general, for standardized outcomes, with mean zero and standard deviation one, we are powered to detect a 0.22 standard deviation change. Specifically looking to our IPV outcomes, we are powered to detect a 3 percentage point reduction in any IPV often (from a baseline rate of 3%) and a 5 percentage point reduction in any IPV in the last 12 months (from a baseline rate of 8%).

We pre-registered the following outcomes:

1. SRH behavior and biomarkers: sexual activity; pregnancies; knowledge and use of contraception methods; HIV and STI knowledge, status, and testing; and family planning goals and strategies;

2. Other health behaviors: intimate partner violence; smoking, drinking, socializing, physical activity, and self-reported physical and mental health;

3. Behavioral economic model parameters: time discounting, risk tolerance, locus of control.

We generate nine indexes across these outcomes. Each index is comprised of several indicators, which are standardized to the mean and standard deviation among females in the control group at baseline and endline separately. The final index is the unweighted mean of the standardized sub-measures. Detail on the index subcomponents is presented in Table D1. We estimate the following specification for all primary outcome measures:

$$
\begin{aligned}
Y_{ict} = \alpha + \beta_1 \text{Boys}_c \times \text{Post}_t + \beta_2 \text{Goal}_i \times \text{Post}_t + \gamma_1 \text{Boys}_c \times \text{Post}_t \times \text{Goal}_i \\
+ \theta_1 \text{Goal}_i + \theta_2 \text{Post}_t + \theta_3 \text{Goal}_i \times \text{Boys}_c + X_{ict}^{'} \xi + \alpha_c + \epsilon_{ict}
\end{aligned} \tag{11}
$$

where Y_{ict} is the outcome of interest for individual i in club c at time t, Supply_c, Boys_c and Goal_i are binary indicators for being assigned to the *Supply*, *Boys*, and *Goal* treatments, respectively, and Post_t is a dummy variable that takes on the value one for the period after treatment is implemented. X_{ict} is a vector of controls including $\text{Supply}_c \times \text{Post}_t$, $\text{Supply}_c \times \text{Goal}_i$ and $\text{Supply}_c \times \text{Post}_t \times \text{Goal}_i$ to control for assignment to the *Supply* treatment and a set of individual characteristics. α_c is a vector of club fixed effects that control for club-level treatment assignment and to account for the stratification of the *Goal* treatment assignment. The standard errors ϵ_{ict} are clustered at the club level to account for the study design. The parameters of interest, β_1 and β_2, capture the ITT effects of the *Boys* treatment and *Goal* treatment, and γ_1 estimate the interaction between the *Boys* treatment and the *Goal* treatment.

Table D2 presents estimates of β_1, β_2, and γ_1 for all outcomes.

Table D1 Registered Outcomes

(1)	(2)
Measure	**Constructon**
Sexual activity index	Comprised of three indicators: (1) a binary indicator that the female has ever had sex; (2) a binary indicator that the female currently has a partner; and (3) an indicator measuring the total number of sexual partners the female reported.
Family planning index	Comprised of four indicators: (1) a binary indicator that the female has never been pregnant; (2) a binary indicator that the female uses injectable contraceptive; (3) a binary indicator that the female uses implant contraceptives; and (4) a binary indicator that the female uses an IUD contraceptive.
IPV often index	Comprised of three indicators: (1) a binary indicator that the female reports experiencing psychological violence often; (2) a binary indicator that the female reports experiencing physical violence often; and (3) a binary indicator that the female reports experiencing forced sex often.
IPV in last year index	Comprised of three indicators: (1) a binary indicator that the female reports experiencing psychological violence in the last year; (2) a binary indicator that the female reports experiencing physical violence in the last year; and (3) a binary indicator that the female reports experiencing forcecd sex in the last year.
Knowledge Index	Comprised of seven indicators: (1) a binary indicator that the female has learned about domestic violence in school; (2) a binary indicator that the female has learned about sexual reproductive health (SRH) in school; (3) a binary indicator that the female has learned about domestic violence or SRH outside of school; (4) a HIV knowledge index comprised of four HIV knowledge indicators (older partners increase risk of HIV, pregnant woman with HIV can give the virus to her unborn baby, condoms reduce risk of HIV, males have higher risk of HIV); (5) an index of contraceptives that the female has heard of (condom, IUD, injectable, pills, implant, female condom, vaginal ring); (6) an index of contraceptives that the female knows how to use (condom, IUD, injectable, pills, implant, female condom, vaginal ring); and (7) an indicator that the female knows of a place to get contraceptive advice.
HIV/STI positive index	Comprised of two indicators: (1) an indicator that the female tested positive for HIV and (2) an indicator that the female tested positive for an STI.
Other health behaviors index	Comprised of three indicators: (1) an indicator that the female does not smoke; (2) an indicator that the female does not drink; and (3) an indicator of the reported hours spent on physical activity on a typical day.
Behavioral index	Comprised of four indicators: (1) a locus of control index comprised of seven statements from Locke and Latham (1990); (2) an indicator that the female is confident she can complete any task she starts; (3) a discount factor to measure patience; (4) an binary indicator that the female is risk averse.
Mental health index	Sum of three items from the PHQ-9 (depressed, low energy, low concentration) standardized to the mean and standard deviation among females in the control group at baseline and endline separately (Kroenke, Spitzer and Williams, 2001). Higher values indicate worse mental health.

Table D2 Impact of Treatments on Registered Outcomes

	(1)	(2)	(3)	(4)
	Boys Treatment	**Goal Treatment**	**Endline Control Mean**	**Observations**
Sexual activity Index				
treatment x post	-0.125**	0.065	0.000	5182
	(0.058)	(0.064)		
treatment x post x Goal	-0.027			
	(0.081)			
Family planning index				
treatment x post	-0.011	-0.042	0.000	5,182
	(0.037)	(0.046)		
treatment x post x Goal	0.047			
	(0.069)			
IPV index				
treatment x post	-0.190**	-0.248**	0.000	5,182
	(0.082)	(0.097)		
treatment x post x Goal	0.110			
	(0.173)			
Knowledge index				
treatment x post	0.107	0.053	0.000	5,182
	(0.086)	(0.054)		
treatment x post x Goal	-0.077			
	(0.072)			
HIV/STI positive index				
treatment x post	-0.023	0.031	0.000	4,122
	(0.079)	(0.101)		
treatment x post x Goal	0.039			
	(0.121)			
Other health behaviors index				
treatment x post	0.216**	0.090	0.000	5,182
	(0.093)	(0.071)		
treatment x post x Goal	-0.036			
	(0.110)			
Behavioral index				
treatment x post	0.038	0.045	0.000	5,182
	(0.075)	(0.059)		
treatment x post x Goal	0.029			
	(0.081)			
Mental health index				
treatment x post	0.007	0.051	0.000	5,182
	(0.142)	(0.079)		
treatment x post x Goal	0.027			
	(0.134)			

Notes. This table presents estimates of β_1 and β_2 and γ_1 from equation 11. For each outcome, the coefficients from a single regression are presented in two sub-rows, with estimates of β_1 in column 1 and β_2 in column 2, labeled treatment x post, and the estimates of γ_1 in column 1 in the second of the two rows, labeled treatment x post x Goal. Column 3 presents the control mean at endline, and column 4 shows the number of observations in the model. All specifications include club fixed effects. See Table D1 for outcome definitions. Standard errors, clustered at the club level, are presented in parentheses below the coefficient estimates in columns 1–2. ***p<.01, **p<.05, *p<.10.
Source. Female baseline and endline data, balanced panel.

E Model Appendix

E.1 Proof: All males with $s_m > 0$ propose

Lemma 4 *It is WLOG to assume that in all equilibria, all males with $s_m > 0$ always propose risky sex.*

Proof. Given that males use violence with probability p upon being rejected, their belief of the female leaving upon receiving violence is

$$q(p) = \frac{\alpha\rho^H(p)}{\alpha\rho^H(p) + (1-\alpha)\rho^L(p)}$$

Suppose $q = 1$ (i.e., all females choose to exit upon receiving violence), then no male will choose violence because it will result in the dissolution of the relationship ($-d_m < 0$), i.e. $p = 0$ (following equation 9). But $q(0) = \frac{\alpha\rho^H(0)}{\alpha\rho^H(0)+(1-\alpha)\rho^L(0)} = \alpha < 1$, i.e. since there is no threat of violence, low-type females will reject risky sex, which results in a decreased exit rate upon receiving violence. Therefore, $q = 1$ cannot appear in the equilibrium.

Now suppose $q < 1$. Then, following equation 9, we know $p > 0$, i.e., since there are female who choose to stay upon receiving violence, then there will always be male who choose to use violence upon receiving no as an answer. With $p > 0$, both ρ^H and ρ^L are less than 1, which means there are always females who say yes to risky sex, so it is strictly better for males with $s_m > 0$ to propose risky sex. To summarize:

Male's payoff from playing "Not propose": 0

Male's payoff from playing "Propose":

$$Pr(Yes)s_m + (1 - Pr(Yes)) \max\{0, qv_m + (1-q)(-d_m)\} > 0$$

E.2 Proof of Proposition 1.

Proof. There exists a unique equilibrium $(p*, q*)$.

This proof has three parts.

(i) We show that $q(p)$ is increasing in p under a simplifying assumption. We define the following elasticity as

$$\varepsilon_p^k = \frac{d\rho^k}{dp}\frac{p}{\rho^k},$$

which measures the relative change in probability the female says no for a change in the conditional probability of violence.

Assumption 1. The elasticity of *low-types* is greater than the elasticity of *high-types.*

$$\frac{|\varepsilon_p^L|}{|\varepsilon_p^H|} > 1 \tag{12}$$

Differentiate (6) with respect to p to get

$$\begin{aligned}\frac{\partial q}{\partial p} &\propto -\alpha f^H d_f^H[\alpha(1-F^H) + (1-\alpha)(1-F^L)] - [-\alpha f^H d_f^H - (1-\alpha) f^L v_f]\alpha(1-F^H) \\ &= \alpha(1-\alpha)\left[v_f f^L(1-F^H) - d_f^H f^H(1-F^L)\right]\end{aligned}$$

Under Assumption 1, we have

$$\frac{|\varepsilon_p^L|}{|\varepsilon_p^H|} = \frac{f(pv_f)pv_f}{1-F(pv_f)} / \frac{f(pd_f^H)pd_f^H}{1-F(pd_f^H)} > 1.$$

Therefore, $\frac{\partial q}{\partial p}$ is positive, so q increases in p.

(ii) Observe that $p(q)$ is decreasing in q. This is an immediate consequence of (9).

(iii) Given (i) and (ii), $p(q)$ and $q(p)$ intersect once, at the equilibrium (p^*, q^*).

E.3 Proof of Proposition 2

Proof. The *Boys* treatment unambiguously reduces violence.

The *Boys* intervention can decrease s_m and/or decrease v_m. If the *Boys* intervention reduces s_m, this reduces violence by reducing the share of males who propose risky sex without affecting the equilibrium (p^*, q^*).

If the *Boys* intervention reduces v_m, this unambiguously reduces p. This has an ambiguous impact on violence. To see this, observe that the probability of violence is P(violence) = (1-α)P(violence|low-type) + αP(violence|high-type), where P(violence|k-type) = P($s_m > 0$) $\times \rho^k(p^*) \times p^*$. The ambiguity is a result of the increase in $\rho^k(p)$ that occurs with a decrease in p. We introduce an intuitive assumption to eliminate this ambiguity.

Assumption 2. The response elasticity of low-type females is less than 1.

$$|\varepsilon_p^L| < 1$$

This assumption bounds the response sensitivity of low-type females, such that the relative increase in probability females say no does not outweigh the decrease in p.

To see how this assumption ensured a decrease in low-type violence, recall from equation 5 that

$$\rho^L(p) \times p = (1 - F(pv_f)) \times p.$$

Then for low-type violence to decrease, we need

$$\frac{\partial(\rho^L(p) \times p)}{\partial p} = 1 - F(pv_f) - f(pv_f)pv_f > 0$$

which is naturally true under Assumption 2. As p decreases, more low-type females will say no since there is less threat of violence. Assumption 2 ensures that this increase in probability of saying no will not outweigh the decrease in the conditional probability of violence.

Further, Assumptions 1 and 2 together imply that $|\varepsilon_p^H| < 1$, which also guarantees that high-type violence will decrease when p decreases. Together, this implies a overall decrease in violence.

E.4 Ambiguity of the *Goal* intervention on violence

In the model, the *Goal* intervention translates into an increase in the net cost of risky sex, s_f, across the distribution of females. Previously, s_f was distributed according to the cumulative distribution function $F(s_f)$. After the intervention, the distribution becomes $F(s_f - \Delta)$, where Δ is equal to the shift of the distribution of s_f.

There are two sources of ambiguity in the impact of the *Goal* intervention on violence. First, the impact of the *Goal* intervention on violence depends on the relative response sensitivity of high-type and low-type females. If low-types' s_f is more affected by the intervention, the conditional exit curve $q(p)$ will shift down and result in a lower q^* and a higher p^*; if high-types' s_f is more affected by the intervention, the conditional exit curve $q(p)$ will shift up and result in a higher q^* and a lower p^*. In the first case, violence increases. In the second case, the impact on violence remains ambiguous despite the decrease in p^*. The result is formally stated in Proposition 3'.

Proposition 3'

Under Assumptions 1 and 2,

1. if $1 < \frac{v_f}{d_f^H} < |\varepsilon_p^L/\varepsilon_p^H|$: the *Goal* intervention will shift $q(p)$ downwards, resulting in a lower q^* and higher p^*, and violence increases;

2. if $1 < |\varepsilon_p^L/\varepsilon_p^H| < \frac{v_f}{d_f^H}$: the *Goal* intervention will shift $q(p)$ upwards, resulting in a higher q^* and lower p^*, and the impact on violence remains ambiguous.

To understand why an increase in q^* and a decrease in p^* is not sufficient to ensure a reduction in violence, notice that the *Goal* intervention increases the female's net cost of risky sex, s_f, and, therefore, increases the probability of females rejecting risky sex at all levels of p. Therefore, although p^* decreases to p', $\rho^k(p)$ increases to $\rho^{k'}(p)$ for all levels of p. Therefore, the overall change remains ambiguous. Assumption 2 ensures that

$\rho^k(p^{'}) \times p^{'} < \rho^k(p^*) \times p^*$. However, it does not ensure that $\rho^{k'}(p^{'}) \times p^{'} < \rho^k(p^*) \times p^*$. This requires a stronger assumption. The proof is as follows:

Proof. After the intervention, the probability of a female exiting upon violence becomes

$$q'(p) = \frac{\alpha(1 - F(pd_f^H - \Delta))}{\alpha(1 - F(pd_f^H - \Delta)) + (1 - \alpha)(1 - F(pv_f - \Delta))}$$

The intervention will shift $q(p)$ upwards if $\partial q'/\partial\Delta > 0$:

$$\alpha f^H[\alpha(1 - F^H) + (1 - \alpha)(1 - F^L)] - [\alpha f^H + (1 - \alpha)f^L]\alpha(1 - F^H) > 0$$

Or equivalently,

$$\frac{f^L}{1 - F^L} < \frac{f^H}{1 - F^H}$$
$$\frac{|\varepsilon_p^L|}{|\varepsilon_p^H|} < \frac{v_f}{d_f^H}$$

Throughout the paper, we always assume that $\frac{|\varepsilon_p^L|}{|\varepsilon_p^H|} > 1$ and we have $v_f/d_f^H > 1$. Therefore, we have the two cases mentioned above:

1. if $1 < \frac{v_f}{d_f^H} < |\varepsilon_p^L/\varepsilon_p^H|$: the intervention will shift $q(p)$ downwards, resulting in a lower q^* and higher p^*;

2. if $1 < |\varepsilon_p^L/\varepsilon_p^H| < \frac{v_f}{d_f^H}$: the intervention will shift $q(p)$ upwards, resulting in a higher q^* and lower p^*.

To reduce the probability of violence against low-type females, $Pr(s_m > 0)(1-\alpha)\rho^L p$, we need that $(1 - F(p'v_f - \Delta))p'$ decreases in Δ, i.e.:

$$\frac{\partial(\rho^L p)}{\partial\Delta} = \frac{\partial p}{\partial\Delta} - F^L\frac{\partial p}{\partial\Delta} + pf^L\left(-v_f\frac{\partial p}{\partial\Delta} + 1\right) < 0$$

Equivalently,

$$\left(1 - F^L - pv_f f^L\right) \frac{\partial p}{\partial \Delta} + pf^L < 0$$
$$\left[\frac{\partial p}{\partial \Delta} - \frac{1}{v_f}\right] |\varepsilon_p^L| > \frac{\partial p}{\partial \Delta}$$

Under case 1, we have $\frac{\partial p}{\partial \Delta} > 0$, so for violence to decrease, we need

$$\left(1 - \frac{1}{v_f \frac{\partial p}{\partial \Delta}}\right) |\varepsilon_p^L| > 1$$

which always fails under Assumption 2. Therefore, under Assumption 2, violence against low-types always increases in case 1: although females will say no more often as p increases, their behavioral response is inelastic; therefore, the increase in p is the dominant effect.

Under case 2, we have $\frac{\partial p}{\partial \Delta} < 0$, so for violence to decrease, we need

$$\left(1 - \frac{1}{v_f \frac{\partial p}{\partial \Delta}}\right) |\varepsilon_p^L| < 1,$$

which is not guaranteed by Assumption 2. The interpretation is as follows: there are two competing forces here determining the overall change in violence. On the one hand, the intervention shifts $q(p)$ upwards, resulting in a lower violent response rate p^*, which reduces the probability of violence; on the other hand, the intervention also increases the share of females who will say no, shifting $\rho^k(p)$ upward to $\rho^{k'}(p)$, which increases the probability of violence. Assumption 2 only guarantees that the change in probability the female says no caused by a change in p will not outweigh the change in p. Therefore, to ensure that the decrease in p outweighs both the shift in and the movement along $\rho^k(p)$, a stronger condition on female's response elasticity is required.