Depression, Violence and Power

Victoria Baranov, Sonia Bhalotra, Pietro Biroli,
Michelle Escobar and Joanna Maselko *

February 12, 2026

Abstract

Does treating maternal depression reduce women's risk of intimate partner violence, or could it sometimes increase it? Harnessing a psychotherapy trial and eight years of follow-up data, we show that treatment reduced maternal depression but increased women's likelihood of experiencing violence. Patterns of heterogeneous treatment effects reveal divergent trajectories for two distinct subgroups of women. Among women with higher status relative to their husbands, psychotherapy reduced sexual violence and increased women's voice: they participate more in household decision-making, have higher labor force participation, and achieve lower fertility. Among lower-status women, similar reductions in depression led to increased psychological, physical and sexual violence and reduced decision-making power, consistent with husbands using violence instrumentally to suppress women's self-advocacy. The results show that the gold standard treatment for maternal depression can have serious adverse effects in the absence of sufficient leverage to support women's restored self-advocacy.

Keywords: psychotherapy, depression, intimate partner violence, power

JEL Codes: J12, J16, D91, I14

*Baranov: Department of Economics, University of Melbourne (email: victoria.baranov@unimelb.edu.au); Bhalotra: Department of Economics, University of Warwick (e-mail: sonia.bhalotra@warwick.ac.uk); Biroli: Department of Economics, University of Bologna (e-mail: pietro.biroli@unibo.it); Escobar Carias: Department of Economics, University of Melbourne (email: michelle.escobarcarias@unimelb.edu.au); Maselko: Gillings School of Global Public Health, University of North Carolina (email: jmaselko@unc.edu). We received helpful feedback from seminar participants at the the University of Pittsburgh, the Monash Centre for Health Economics and the University of Tasmania, and conference participants at NEUDC, PacDev, the Australasian Meeting of the Econometric Society, AGEW, ADEW, and ADEA-QUT.

1 Introduction

Does treating maternal depression reduce women's exposure to intimate partner violence—or could it sometimes increase it? This question has high stakes: maternal depression and intimate partner violence (IPV) are among the leading contributors to the global disease burden (Vos et al., 2012), frequently co-occurring and affecting hundreds of millions of women worldwide. Depression is characterized by diminished self-advocacy, assertiveness, and ability to articulate preferences (Beck 1976; Locke et al., 2018; Leykin et al., 2011), and can generate conflict due to breakdowns in communication and emotional withdrawal. One might expect, then, that treating depression could reduce violence by improving women's self-advocacy and spousal communication.

Yet, economic models of the household predict that this may not always be the case. Husbands may use violence for different purposes: expressive (for its own sake or to release frustration) or instrumental (to extract resources or compliance). When the primary motive is expressive, husbands may use less violence as psychological vulnerability decreases. By contrast, when the primary motive is instrumental, improvements in women's mental health, self-advocacy and resistance to control may be countered by husbands with increased violence. This paper provides the first experimental evidence that treating depression can indeed increase women's exposure to violence.

We exploit a randomized intervention that treated maternal depression to study its effects on women's exposure to multiple forms of violence. The intervention was based on the principles of cognitive behavioral therapy, the clinical gold standard for treating mild to moderate depression. It targeted prenatally depressed women in rural Pakistan and was delivered by trained peer volunteers from women's third trimester of pregnancy until the focal child reached age five (Sikander et al., 2019; Maselko et al., 2020). Following randomization, women were surveyed at frequent intervals for the next eight years, making this the ideal set up for observing the dynamic relationship between depression and violence.

Rates of IPV victimization are strikingly high in this population. At baseline, 65.6 percent of prenatally depressed women reported experiencing some form of IPV in the past 12 months—compared to 35 percent among women who were not prenatally depressed. Moreover, baseline maternal depression severity is strongly and positively correlated with a higher frequency of violence. Similar patterns have been documented in the public health literature, where poor mental health is considered an important risk factor for IPV victimization (Devries et al., 2013).

Our experimental results reveal a striking pattern that challenges this view: despite being highly effective at reducing depression, treatment led on average to a modest in-

crease in intimate partner violence. Specifically, women who were assigned to receiving psychotherapy experienced large and persistent reductions in depression, averaging an 8.2 percentage point decline over the eight years following treatment, yet also faced a 4.2 percentage point increase in violence, with both physical and psychological IPV rising. Leveraging data from a list randomization experiment, we show that these increases reflect higher incidence of violence rather than changes in reporting behavior. These average treatment effects are consistent with an instrumental motive for violence dominating, and with husbands using violence to suppress women's self-advocacy.

Economic models of the household predict that these effects should depend not only on husbands' motives for violence, but also on women's leverage within the relationship (Tauchen et al., 1991; Farmer and Tiefenthaler, 1997). Women who possess greater leverage may be better positioned to resist violence and credibly threaten to withdraw cooperation or exit the relationship.[1] We test this prediction by examining treatment effect heterogeneity by women's status relative to their husbands.

We proxy for women's status using the gap in schooling years between spouses. We validate this proxy by showing that women with more schooling than their husbands (approximately one-third of the sample) exhibit greater influence over household decision-making and a higher ability to realize their fertility preferences when spouses disagree over desired family size. Because schooling itself is a productive attribute that may independently affect outcomes, we condition on women's absolute level of education throughout our analysis. Our main results focus on women with secondary education or higher, the subgroup for whom the intervention was effective at reducing depression.

Among higher status women, treatment leads to large reductions in sexual violence (8.7 percentage points), a form of abuse closely linked to expressive motives (Haushofer et al., 2019; Tarzia, 2021). By contrast, among women with lower status, reductions in depression are followed by large increases in psychological (9.6 percentage points), physical (12.1 percentage points) and sexual violence (11.1 percentage points). The former two are more closely associated with strategic control and instrumental motives (Haushofer et al., 2019).

These patterns extend beyond IPV to other, more cooperative domains of household bargaining. Among women with higher status, treatment is accompanied by increased involvement in household decisions (0.28 SD), higher labor force participation (6.2 percentage points), and reductions in cumulative fertility that align with women's stated preferences,[2] consistent with women advocating more for outcomes that matter

[1] In practice, women in rural Pakistan are more likely to threaten her husband with going to stay with her parents, rather than seek a divorce.

[2] This result is consistent with the hypothesis that when spouses diverge over fertility preferences, women's bargaining position plays an important role in shaping reproductive outcomes (Doepke and

to them and being able to realize their goals. By contrast, among women with lower status, treatment leads to reduced involvement in household decision-making (0.209 SD) and lower labor force participation (6.9 percentage points). This is consistent with women self-censoring to avoid further violence, with men using these domains as additional tools to penalize women for their self-advocacy, or with both mechanisms operating simultaneously. Together, these results indicate that the IPV effects are not isolated, but instead reflect broader changes in household bargaining dynamics following reductions in women's depression.

This paper makes two primary contributions to the literature. First, we contribute to a growing public health literature studying the bi-directional relationship between intimate partner violence and mental health disorders (Trevillion et al., 2012; Oram et al., 2022). To the best of our knowledge, this is the first study to document adverse effects of depression treatment for current and potential IPV survivors on future victimization. The closest related interventions are trauma-informed therapies explicitly designed for IPV survivors, which combine mental health treatment with psychoeducation on recognizing abuse patterns and safety planning (Tol et al., 2019; Spencer et al., 2019). In contrast, we study a general psychotherapy intervention for maternal depression that does not include IPV-specific content, allowing us to isolate the effects of treating mental health symptoms themselves rather than trauma-informed safety interventions.

Second, we contribute to the literature on the economics of intimate partner violence and empowerment. A large economics literature documents that interventions aimed at improving women's economic position can have heterogeneous effects on IPV. Some studies find that improvements in women's outside options reduce violence, including divorce law reforms (Stevenson and Wolfers, 2006), income gains and transfers (Angelucci, 2008; Bobonis et al., 2013; Hidrobo et al., 2016), and reductions in gender wage gaps (Aizer, 2010). Other studies document backlash effects, where interventions that shift women's roles, opportunities, or norms increase violence or provoke other forms of intrahousehold resistance in particular contexts or subgroups (Roy et al., 2018; Erten and Keskin, 2018, 2021; Guarnieri and Rainer, 2021; Cullen, 2020; Abou Daher et al., 2025). Standard economic models rationalize these mixed findings by allowing violence to serve different functions (Tauchen et al., 1991; Farmer and Tiefenthaler, 1997; Eswaran and Malhotra, 2011; Card and Dahl, 2011).

We build on this literature by incorporating the role of mental health in shaping self-advocacy and preference articulation. In this framework, improving one partner's mental health alters intrahousehold interactions by changing how preferences are expressed without mechanically changing underlying bargaining positions. This per-

Tertilt, 2018).

spective highlights the complementarity between mental health and leverage: leverage without the capacity for advocacy may remain latent, while increased advocacy in the absence of leverage can provoke conflict. More broadly, this insight helps explain heterogeneous responses to empowerment-oriented interventions and extends beyond violence and households.

The rest of the paper proceeds as follows. Section 2 outlines a conceptual framework linking treatment for depression to intimate partner violence. Section 3 describes the study context and the experimental design. Section 4 details the data employed in the analyses and Section 5 describes our empirical strategy. Section 6 presents average treatment effects on depression and IPV. Section 7 focuses on the role of women's status. Section 8 explores alternative mechanisms and child outcomes, and Section 9 concludes.

2 Conceptual Framework

Violence within the household is inherently a two-agent outcome, shaped jointly by women's outside options and husbands' motives for violence. We begin by distinguishing between two primary motives for violence: expressive and instrumental.

Expressive violence is taste-based, reflecting men's preferences or predispositions toward violence, such that IPV enters the husband's utility function directly and independently of its effect on material allocations or household decisions (Tauchen et al., 1991). Empirical work links such violence to entitlement and emotional cues, rather than strategic control (Card and Dahl, 2011; Haushofer et al., 2019; Tarzia, 2021). When violence is primarily expressive, perpetrators may be more likely to inflict abuse when women are psychologically vulnerable and less able to resist or mobilize social and familial support. We hypothesize that in this setting, depression may increase women's exposure to IPV, and improvements in mental health could reduce violence by strengthening women's ability to articulate preferences, advocate for their goals, and seek help.

Instrumental violence, by contrast, is strategic. It is used as a bargaining tool to influence behavior, extract compliance, and maintain control within the household (Eswaran and Malhotra, 2011; Tauchen et al., 1991). In such models, violence is employed not because it is valued per se, but rather to alter the bargaining outcome (Bloch and Rao, 2002). We therefore posit that when violence is primarily instrumental, improvements in mental health may instead backfire as the increases in self-advocacy that accompany reductions in depression may raise conflict and trigger violence aimed at suppressing that advocacy. Classical bargaining theory also predicts that such violence is more likely when women have limited leverage within the household, since violence

is less costly when women lack credible threat points such as the ability to exit the relationship or withdraw cooperation (Farmer and Tiefenthaler, 1997; Lundberg and Pollak, 1993; Manser and Brown, 1980; McElroy and Horney, 1981).

This framework yields two clear, testable predictions. First, improvements in women's mental health may reduce expressive forms of violence, particularly where women have sufficient leverage to act on increased self-advocacy. Second, the same improvements may instead be followed by increases in instrumental forms of violence among women who lack leverage within the household, as a means to curtail self-advocacy. As a result, improvements in mental health may reduce some forms of IPV while increasing others, depending on women's bargaining position within the household.

Building on this conceptual framework, the following section outlines the experimental design used to evaluate whether treatment for depression increases or reduces IPV.

3 Experimental Design

The Thinking Healthy Program, Peer-delivered Plus (THPP+) was a cluster Randomized Controlled Trial (RCT) addressing perinatal depression in rural Pakistan (Maselko et al., 2020). Kallar Syedan, the rural subdistrict of Rawalpindi, Punjab, chosen for the intervention is a socioeconomically deprived area with a poverty rate of about 50%, a high prevalence of maternal depression and little access to clinical mental health care (NIPS & ICF, 2013).

Randomization.— The trial was randomized across 40 village clusters: 20 clusters were assigned to receive the intervention and 20 to the control arm. The clusters were geographically separate to minimise contamination risk. Eligible participants were women aged 18 years or older, in their 3rd trimester of pregnancy, married and registered with a village-based health worker (also known as Lady Health Workers or LHWs) (Sikander, Ahmad, Bates, et al., 2019).

Between October 2014 and February 2016, eligible women were approached and screened for depression using the Urdu version of the Patient Health Questionnaire (PHQ-9). The tool includes questions about the frequency of depressive symptoms in the last two weeks such as ability to concentrate, sleeping or eating problems and suicidal thoughts. It has been extensively used in the study setting and has an acceptable criterion validity and reliability for this population (Gallis et al., 2018). Women who scored 10 or greater on the PHQ-9 tool, screened positive for depression, and were invited to participate in the trial (Sikander, Ahmad, Atif, et al., 2019).

Out of the 1731 women screened, 572 (33%) were identified as depressed. Among

them, 287 mothers were in clusters randomized to the intervention, 283 in the control arm, with only two mothers refusing to participate before the baseline assessment. Of the 1159 pregnant women who were screened as non-depressed, 584 were randomly selected to be part of the non-depressed arm of the study. They would represent a natural reference group to understand the evolution of maternal and child outcomes.

The THPP+ Intervention.— The Thinking Healthy Program, Peer-delivered Plus (THPP+) is a low-intensity psychosocial intervention delivered by non-specialist, volunteer peer women from the same community as the mother. Peer volunteers were trained in a 5-day classroom-based workshop, followed by a 2-month internship during which they practiced the content of the intervention on non-trial participants. They also received monthly supervision throughout the trial (Turner et al., 2016).

THPP+ consisted of ten individual and four group sessions delivered from the third trimester of pregnancy until six months after child-birth. From seven months to 36 months postnatal, 18 group-based booster sessions were delivered. The first six were delivered monthly, and the rest every two months. The first 14 sessions focused on identifying unhealthy behaviors and inducing behavioral activation. The boosters focused on self-care and infant development by encouraging mother-infant interaction and play. The group sessions also provided a safe environment for women to voice their problems, share experiences and support each other.

Perinatally depressed women in the treatment arm received the THPP+ intervention. In addition to this, both intervention and control groups received enhanced usual care (EUC). The latter consisted of informing participants about their depression status and ways to seek help for it during and after pregnancy. Women without prenatal depression were not offered treatment.

4 Data

We provide a brief description of the data and outcomes below.

Sample and Follow-up.— Our sample consists of the experimental groups of depressed mothers in the treatment and control arms. Data collection on the mother-child dyads has taken place 11 times since the trial (3rd trimester of pregnancy, 3, 6, 12, 24, 36, 48, 72, 73, 84, and 96 months postnatal), as shown in Figure 2. However, not all waves have data on the experience and intensity of intimate partner violence. The effectiveness of the intervention on maternal depression and child development at three years postnatal has been evaluated in previous studies (Maselko et al., 2020; Sevim et al., 2024).

Baseline Balance.— The experimental sample was slightly imbalanced at baseline, as shown in Table 1. For example, women in the treatment arm were approximately 1.1 cm taller, had a smaller waist circumference by 1.79 cm, and lived in households with 0.37 fewer people. They also suffered from slightly higher, albeit statistically insignificant, depression scores, as well as psychological, physical, and sexual IPV. A joint F-test confirms the balance of baseline characteristics (p-value = 0.09).[3]

Attrition.— In Figure 2 we show lost to follow-up rates over time. By year 7 post birth of the index child, lost to follow-up was 25.7% and by year 8, 66.6% of the original 1,154 women remained in the sample. Most attrition took place during the 3-month postpartum wave. The main reason for this behaviour was the loss of the index child (40% of the attritors), which was balanced across study arms. In Appendix Table B3, we show that jointly testing for the difference in attritor characteristics by treatment arms yields balance in all follow-up waves, except for year 8 (p-value = 0.01).

A. Primary Outcomes

Mental Health.— To measure maternal mental health, we use the Patient Health Questionnaire (PHQ-9) and the Structured Clinical Interview for DSM (SCID). The PHQ-9 is a 9-item scale used to assess severity of depression symptoms in the last two weeks and it ranges from 0-27 (Kroenke et al., 2001). Response options include "not at all", "several days", "more than half the days" and "nearly every day". A higher score is indicative of greater severity of symptoms of depression. PHQ-9 scores were measured in all waves except at year 1 of follow-up. Additionally, we use the SCID (Gorman et al., 2004) to construct an objective measure of depression status across all waves, including year 1. At baseline, 73% of women in the control arm were diagnosed with a current major depressive disorder, as shown in Table 1.

Intimate Partner Violence.— In years 1, 2, 3, 6 and 7 of follow-up, we measured the experience of intimate partner violence (psychological, physical and sexual) in the past 12 months using a modified version of the WHO Violence Against Women Instrument (WHO, 2005).[4] Additionally, in year 8 of follow-up, we measured the experience of physical violence. Respondents were also asked about the frequency of each act in the last 12 months. Response options included "1-2 times", "3-5 times", "6-10 times", "11-

[3] Appendix Table B2 shows corresponding tests for balance in baseline characteristics across follow-up waves. The samples observed in years 7 and 8 are imbalanced in terms of baseline characteristics. Appendix Tables **??-??** show balance tests across all follow-up waves by schooling gap group.

[4] See Appendix Table B1 for a complete list of all the items.

20 times" and "more than 20 times". Severity of each form of IPV was then calculated in two steps. First, by recoding frequency categories to the value of the midpoint and second, by summing the frequencies of each act. On average, 49% of control women had experienced psychological IPV at baseline, 23% had experienced physical IPV and 37% had experienced sexual IPV.

With 45 IPV-related questions (ever, any and frequency in the past 12 months), this module took a significant amount of time and space in each survey. Given the sensitive nature of the IPV questions, interviewers were trained to ensure that the environment was safe for disclosure, ensuring that the husband and other families were not within hearing distance, and informed respondents that they could skip questions if necessary. All interviewers were female.

B. Secondary Outcomes

Fertility.— Using child rosters and questions on current pregnancy status collected at various points in time, we reconstructed women's fertility patterns across all waves. We also measure desired fertility of the woman and her husband (as reported by the woman).

Decision-making and Labor Force Participation. — In years 1 and 3 of follow-up, we measure female participation in household decision-making through the questions such as: "who makes decisions about healthcare for yourself" and "who makes decisions about major (and minor) household purchases. Response options included: (0) the woman does not participate in decision-making, (1) the woman participates in decision-making with others, and (2) the woman makes decisions alone. In years 1, 3, and 6 of follow-up, we also asked women about decisions regarding their employment outside the household and their role in childrearing. All dimensions are dichotomized to capture whether or not women are involved in household decision-making and then used to create a standardized index of decision-making power as well as a binary variable, which takes a value of one if a woman is involved in all domains and zero otherwise.

5 Empirical Strategy

First, to estimate treatment effects over the eight years of follow-up, which we refer to as treatment effect *trajectories*, we pool all waves of data, including baseline, and estimate the impact of treatment using one regression in a difference-in-differences speci-

fication.[5] First we report the impact of treatment at each follow-up wave by interacting the treatment indicator with an indicator for the wave:

$$Y_{ict} = \alpha + \sum_{j=1}^{7} \beta_j (Treat_c \times \mathbf{1}[t = j]) + \eta Treat_c + \gamma_t + \mathbf{X'_{ic}\Gamma} + \varepsilon_{ict} \quad (1)$$

where Y_{ict} is the outcome for mother i living in cluster c, at time t, where t ranges from 0 (baseline) to up to 8 years post start of the intervention. The variable $Treat_c$ is a dummy equal to 1 if the mother is in the treatment group, which by the cluster design varies only at the village level, c. $\mathbf{1}[t = j]$ are indicators for each post-wave t available (the specific waves that can be included depend on the outcome being measured). γ_t are fixed effects for post-treatment waves. With this specification, baseline measures appear on the left-hand side of the equation, and chance baseline differences between treatment and control groups are controlled for by including the un-interacted $Treat_c$ term.

$\mathbf{X}$ is a vector of characteristics that were imbalanced at baseline (mother's height, waist circumference, number of people per room, asset wealth, frequency of sexual IPV), the baseline level of the outcome when available, and union council fixed effects, all interacted with wave indicators.[6] For outcomes not measured at baseline, analyses are performed separately for each wave. The main parameters of interest, β_j, are the intention-to-treat impacts of being assigned to the intervention on outcomes in wave j.

Second, to estimate the average effect of the intervention across all of the waves, β, we replace the dummies for individual waves in the interaction terms with one single indicator for post wave, $Post_t$, while keeping all other controls identical to specification (1). Specifically, we estimate:

$$Y_{ict} = \alpha + \beta (Treat_c \times Post_t) + \eta Treat_c + \gamma_t + \mathbf{X'_{ic}\Gamma} + \varepsilon_{ict} \quad (2)$$

In all specifications, we account for attrition using inverse probability weights determined by predicting attrition status using baseline characteristics. We report standard errors clustered at the unit of randomization.

[5] This approach can provide more power when autocorrelation is high, as in our case, see McKenzie (2012).

[6] This specification is equivalent to performing ANCOVA regressions but with all waves of data simultaneously, as opposed to wave by wave.

6 Treatment Effects on Depression and IPV

In this section, we present the estimates of the evolution of maternal mental health and intimate partner violence trajectories over eight years after the birth of the index child. Next, we examine how these effects were moderated by women's household bargaining power and other baseline characteristics. The results on mental health build on Sevim et al. (2024), who show that by the time the index child is three years old, maternal mental health and daily functioning were significantly higher among women in the treatment arm.

6.1 Depression Trajectories

Averaging all eight years of follow-up data shows that the intervention was effective in reducing depression across all measures we consider. PHQ-9 depression scores were significantly lower among women who received psychotherapy by 1.117 units (16.5% relative to the control mean), as shown in Table 2. The temporal pattern of effects shows a rapid and sustained recovery (Figure A1), with greater effects at the 3-month followup. By three months postnatal, six months after the start of the intervention, depression severity was lower by 34%; by six months postnatal, PHQ-9 scores were still lower by 18.7% among the treatment arm (see Appendix Table B4). Longer-term results are still remarkably strong: treatment effects on depression scores continued to range between 16% and 19% throughout the eight years of follow-up, with dips at years 2 and 6 (which coincided with COVID-19). After the COVID-19 wave, the intervention effects rebounded, with treated women having 17.5% lower PHQ-9 scores by the eight year of follow-up.

To capture whether these improvements in depression scores were sufficiently large to place patients under clinical depression thresholds, we look at binary indicators of depression. We dichotomize PHQ-9 scores, with values above 10 being indicative of moderate to major depressive disorders (Table 2) and also use the Structured Clinical Interview for DSM (SCID) to obtain an objective measure of clinical depression as assessed by a trained clinician (Table 2). Estimation results using all follow-up periods show that the psychotherapy intervention leads to a significant reduction in the probability of being diagnosed with clinical depression across both instruments, with effects ranging from 5.5 (28%) to 6.7 (23.7%) percentage points for the SCID and PHQ-9 instruments, respectively.

6.2 IPV Trajectories

An index measure capturing the experience of any form of intimate partner violence in the last 12 months is significantly higher among women in the treatment arm, by 4.2 percentage points, or 8.9% relative to the control mean (Table 2). Increases in psychological and physical violence are the main contributors, each one spiking on average by 4.1 (10%) and 3.5 (25.4%) percentage points, respectively, across the entire follow-up period.[7] These treatment effects appear to be concentrated in the medium (2-3 years) and longer term (7-8 years), as depicted in Figure A1 (see Appendix Table B5).

To explore whether violence increased also at the intensive margin, we ask women about the frequency with which they experienced various forms of psychological, physical, and sexual violence in the last 12 months. When averaging across all follow-up periods, the frequency of IPV increases only slightly: the estimated coefficients are small and noisily estimated (Table 2). However, at year 2, there is a significant increase in psychological and physical violence, and at year 7, the frequency of psychological violence spikes again among the treatment arm (Appendix Figure A1).

Since our measures of intimate partner violence are self-reported, there are two potential areas of concern: social desirability bias, which would lead to under-reporting of IPV among the treatment arm, and changes in reporting induced by depression treatment. The latter is particularly concerning, since improvements in mental health and social interactions with the peer volunteer and other depressed patients may increase awareness and understanding of IPV. This may lead to increased reporting of violence among the treatment arm, such that we may wrongly conclude that the intervention increases the incidence of violence.

To address these concerns, we refer to the findings by Bates et al. (2025)[8], which presents the results from a list randomization experiment conducted in year 6 of follow-up among our sample. In this wave, the authors cross-randomized participants to receive either a list of four non-sensitive items or the same list with an additional item about physical violence.[9] After comparing the average number of items reported across groups separately by treatment arm, and benchmarking these figures against direct questions on physical violence, the authors show that: a) there are no differences in reporting among the control group, and b) women in the treatment arm report four

[7] This is comparable in magnitude to effects documented for several large household shocks studied in prior work. For instance, cash transfers to women equivalent to a 35 percent increase in household income reduce aggressive behavior by about 21 percent (Angelucci, 2008), and the introduction of unilateral divorce laws reduces domestic violence by roughly 30 percent (Stevenson and Wolfers, 2006; Brassiolo et al., 2016). The estimated increase is somewhat smaller, though of similar order, to the 32 percent rise in domestic violence observed following male job loss (Bhalotra et al., 2025).

[8] The team is working on this paper and it should be available as a WP soon

[9] "You were slapped, pushed, shoved, or punched by your husband or he threw something at you that could hurt you."

times as much physical violence in the list randomization compared to the control, and almost twice as much physical violence relative to when they are asked directly. These results indicate that if anything, the estimates in Table 2 may be underestimating the treatment-induced increases in intimate partner violence.

7 The Role of Women's Status

In what follows, we first provide descriptive evidence to show that the spousal schooling gap meaningfully proxies women's status in the household (vis-a-vis her husband), and then proceed to show treatment effects on primary and secondary outcomes separately for women with more schooling than their husbands ($W > H$) and those with equal or lower schooling ($W \leq H$) while holding women's absolute schooling fixed.

Validating the Spousal Schooling Gap.— Direct measures of bargaining power (or potential) are often difficult to observe. To create a proxy for this latent construct, we borrow evidence from the fields of sociology and demography which, over the past two decades, have begun documenting the rise of a phenomenon known as hypogamy, the practice of marrying a partner of lower social status. We focus specifically on women's educational hypogamy measured as having at least one more schooling year than one's husband. Approximately 31.6% of women in our sample fall within this category, 23.9% have the same level of educational attainment as their husbands, and 44.5% have lower schooling than their spouses, as shown in Figure 3. To assess the validity of this proxy, we conduct three empirical exercises that demonstrate how this gap is associated with underlying differences in women's power and agency within the household.

First, we show that at baseline, women with more schooling than their husband ($W > H$) were younger and had fewer living children compared to women with equal or lower schooling than their spouses ($W \leq H$). They were also 15 percentage points less likely to be among the poorest half of the asset distribution and 46 percentage points more likely to have attained secondary schooling or higher; their husbands, in contrast, were 24 percentage points less likely to have have secondary schooling or more (see Appendix Table **??**). Using data from the 2017-2018 Pakistani Demographic Health Survey (DHS), we show that our sample of respondents closely resembles women living in rural areas throughout Pakistan and that the patterns described above by schooling gap are consistent across samples.

Next, we examine the association between the educational gap and women's participation in household decision-making. In Panel A of Appendix Table B8, we regress various indicators of decision-making on a binary indicator for $W > H$ couples, con-

trolling for age, number of children, and asset wealth. Results for the DHS sample show that having more schooling than one's husband is strongly and positively correlated with women's involvement in household decision-making. In Panel B, we perform the same analyses for the Bachpan sample and observe strikingly similar patterns. In Panel C, we then show that not only is the spousal schooling gap a strong predictor of female agency, it also dominates the effects of women's and men's schooling when included in the same regression. Lastly, in Panel D, we explore whether the relationship between the schooling gap and agency is moderated by depression. The estimates suggest that depression is associated with lower agency among women with less bargaining power but the loss in agency associated with depression is larger for women who have more schooling than their husbands.

Lastly, we relate the schooling gap with the evolution of women's desired and realized fertility by year 8 (the last year of follow-up), using data from women in the control group and the non-depressed cohort. In Appendix Table B9, we show that women's stated fertility preferences in earlier follow-up waves are predictive of the number of living children by year 8. However, among $W > H$ couples, whether the husband wants more children than the wife is not associated with higher realized fertility. In contrast, in $W \leq H$ couples, husbands' preferences for additional children are positively associated with realized fertility, suggesting that when spousal preferences diverge, women with more schooling than their spouses are more likely to achieve their preferred outcome than women with equal or lower schooling.

Together, these empirical patterns demonstrate that women with more schooling than their husbands are systematically different from those with equal or lower schooling: they have different demographic characteristics, display greater involvement in household decision making at baseline, and are more likely to realize their fertility intentions even in the absence of psychological treatment. These patterns serve as strong motivation for employing the spousal schooling gap as a meaningful, although imperfect proxy for bargaining power.[10]

Mental Health.— We now proceed to show treatment effects on primary outcomes separately by the relative spousal schooling gap, beginning with mental health. In line with the results from the random forests, we find striking differences in treatment effects on depression by baseline schooling gap, as shown in Table 3 and Figure A3. Improvements in mental health are largely concentrated among women who had more

[10]Note that having greater schooling than one's husband may capture more than the woman's bargaining power potential. It may also capture other unobservable traits such as having a more progressive husband or a woman's willingness to marry someone from a lower social status. It may also have different implications in regards to her relationship with her mother-in-law and other household members.

schooling than their husbands. For this group ($W > H$), treatment reduces PHQ-9 depression scores by 2.394 points, representing a 32% decrease relative to the control mean. In contrast, for women with equal or lower schooling than their husbands ($W \leq H$), the reduction is only 0.524 points (7.6%), with the difference between groups being statistically significant ($p = 0.003$).

The same pattern holds for clinical depression thresholds. Among $W > H$ women, treatment reduces the probability of scoring above the PHQ-9 cutoff (>10) by 16.4 percentage points, a 51% reduction relative to the control mean, while the effect among $W \leq H$ women is small and statistically insignificant (-2.4 percentage points). The difference is highly significant ($p \leq 0.001$). Objective diagnoses using the Structured Clinical Interview for DSM Disorders (SCID) confirm this heterogeneity: treatment reduces SCID-based clinical depression by 14.7 percentage points for $W > H$ women, with no significant effect for $W \leq H$ women. These results confirm that the intervention was substantially more effective at improving mental health among women with greater baseline bargaining power.[11] Note as well that when we race the interaction terms for the spousal schooling gap, women's schooling, and her husband's schooling with the treatment indicator, the schooling gap dominates (Appendix Table **??**, panel A), which supports the conclusion that the effects in Table 3 are driven by the spousal gap rather than each person's own schooling level.

Intimate Partner Violence.— In regards to IPV, treatment effects differ by spousal schooling gap, not only in magnitude but also in direction, in line with the results from the causal forest. Table 4 shows that, among women with more schooling than their husbands ($W > H$), treatment *reduces* the likelihood of experiencing sexual violence by 12.5 percentage points (43%), and lowers its frequency by 1.10 occurrences (50%). We also observe negative treatment effects on the probability of physical IPV, as well as the frequency of psychological and physical IPV, though these are imprecisely estimated.

In contrast, for women with lower bargaining potential ($W \leq H$), the intervention *increases* exposure to violence. The probability of experiencing physical violence rises on average, by 7 percentage points (58% relative to the control mean). We also observe an escalation in the frequency of psychological violence by 19%. The differences in treatment effects across the two groups are statistically significant for both the probability of physical and sexual IPV ($p = 0.055$ and $p = 0.005$, respectively).[12] Taken together, these results suggest that the intervention reduced exposure to physical and sexual violence among women with higher baseline bargaining potential,

[11] In Appendix Table **??** we display depression trajectories by schooling gap across all follow-up waves.

[12] In Appendix Table **??** we display IPV trajectories by schooling gap across all follow-up waves, and in Appendix Table B12 we display the treatment effects on the items within each type of violence.

while potentially triggering backlash among women starting in weaker bargaining positions. Importantly, these results also survive after racing the spousal schooling gap, the woman's absolute schooling and her husband's schooling (Appendix Table **??**, panel B). All this evidence is further supported when focusing on "softer" measures of spousal relationship quality which are likely less prone to social desirability and reporting bias. In Appendix Table **??**, we show that treated women with lower baseline bargaining power are less likely to report that their husbands look after their basic needs and understand their feelings, whereas treated women with higher baseline bargaining potential are more likely to report that their relationship quality is good or very good.

Fertility.— Another tangible measure of women's agency is their ability to realize their fertility goals. Fertility outcomes are particularly informative in our context because, compared to self-reports of violence or relationship quality, they are less susceptible to reporting or recall bias. Panel A of Table 5 presents treatment effects two fertility-related outcomes: the probability of being pregnant during the intervention period and the number of additional children born by year 8 of follow-up since the birth of the index child. These outcomes offer a cumulative, behavioral reflection of household decision-making and women's ability to act on their reproductive preferences.

Table 5 reveals substantial differences in fertility-related outcomes by baseline bargaining power. Throughout the follow-up period, treated women with greater bargaining potential ($W > H$) are 4 percentage points less likely to be pregnant at any given wave. By year 8 of follow-up, women in this group have 0.466 fewer children born after the index child. These effects are large, amounting to a 36.4% reduction in cumulative fertility over the follow-up period relative to the control group. In contrast, among women with equal or lower schooling than their husbands ($W \leq H$), treatment has no statistically significant effect on women's cumulative fertility or likelihood of being pregnant.

Decision-Making and Labor Force Participation.— Along with fertility, measuring women's involvement in household decision-making allows us to provide a more comprehensive view of cooperative aspects of household bargaining. In Panel B of Table 5, we show that among women with greater baseline bargaining potential (W > H), assignment to treatment increased an index of women's decision-making power by 0.24 standard deviations and increased women's likelihood of being involved in **all** domains of household decision-making (78% increase relative to the control mean). As with fertility, we see no statistically significant treatment effects on decision-making

for women with lower bargaining potential (W $\leq$ H).

For women with higher status (W $>$ H), treatment also led to greater labor force participation by 12.8 percentage points. In contrast, for women with lower status (W $\leq$ H), treatment reduced labor force participation by 6.9 percentage points, compared to the control group.

8 Alternative Mechanisms and Child Outcomes

To better understand the mechanisms that may contribute to the observed backlash among women with lower bargaining potential, we examine exploratory evidence on intermediate outcomes, including social capital, gender norms, time and risk preferences, and parenting and child development. These outcomes reflect channels through which improved mental health could plausibly affect household conflict. For example, increases in social capital may alter women's support networks and facilitate the transmission of new norms and attitudes. In turn, shifts in gender norms may influence how women negotiate their roles within the household. On the other hand, changes in time and risk preferences may affect intertemporal decisions that require cooperation or agreement with partners, and improvements in parenting or child behavior may modify sources of stress or conflict in daily life.

8.1 Alternative Mechanisms

Social Capital.— Between the third trimester of pregnancy and the third birthday of the index child, women in the treatment arm were invited to participate in 22 group sessions, providing a structured context for repeated peer interaction and shared problem-solving. These features may have fostered greater social connectedness. Panel A of Table 6 reports treatment effects on women's social interactions by the spousal schooling gap. In both schooling-gap groups, treatment increased interactions with other pregnant women, an effect that likely arises partly mechanically from group participation, but it also meaningfully enhanced perceived social support from friends and women's willingness to discuss experiences of intimate partner violence with friends or relatives.

Gender Norms.— These group interactions may have also increased exposure to alternative role models and discourses around women's roles, potentially fostering shifts in gender norms change. Panel B of Table 6 shows that by years 7–8 of follow-up, the intervention produced more progressive gender norms on a standardized index, with

treatment effects of roughly 0.2 standard deviations in both schooling gap groups. The index captures attitudes toward female labor force participation as well as the attribution of traits such as bravery and competitiveness to men and women alike.

Time and Risk Preferences.— Across both spousal schooling gap groups, we find no evidence that the intervention affected time or risk preferences. Point estimates are small and statistically insignificant, as shown in Panel C of Table 6, suggesting that the program did not shift these deeper preference parameters. These results indicate that the economic and household behavior changes documented earlier are unlikely to be driven by shifts in risk tolerance or patience.

Parental Investments.— Across

8.2 Child Outcomes

Lastly, we examine whether changes in child development could account for the divergent IPV trajectories observed across household types. Panel D of Table 6 shows a mixed pattern of results. For boys, the intervention has no effect on literacy and small negative effects on numeracy, while for girls we observe positive effects, particularly among mothers with greater bargaining potential. However, these patterns do not map onto the treatment effects on IPV: improvements in child outcomes are neither concentrated in the households where IPV falls nor absent in the households where IPV rises. Taken together, these findings suggest that changes in child development are unlikely to be a mechanism behind the differential IPV responses to reductions in maternal depression.

9 Discussion

This study shows that improving mental health is not merely a therapeutic goal but a pathway to psychological empowerment. Treating depression can restore women's intrinsic agency: their sense of self-worth, motivation, and capacity to envision alternative possibilities. Yet translating this intrinsic agency into external empowerment depends on power structures within the household. In households where women possess sufficient leverage, such psychological gains can have positive effects across both the conflict and cooperative domains of bargaining. Absent that leverage, the same rise in agency may collide with entrenched hierarchies, generating resistance and, in some cases, backlash.

From a policy perspective, these findings caution against scaling psychosocial interventions in isolation. Their success depends not only on what is delivered but also on who delivers it and where it lands. Peer-delivered psychotherapy can be a scalable and inclusive model, yet fidelity and motivation among facilitators are crucial for safeguarding against harm. More broadly, interventions that combine psychological empowerment with changes in opportunity structures, such as skill building, household dialogue, or community-level norm change, may be more likely to yield durable and equitable gains. In settings where women's power is constrained by patriarchal hierarchies, complementing women's therapy with engagement of husbands, mothers-in-law, or other key family members could help align household expectations and sustain psychological progress. Such approaches acknowledge that empowerment is not only an individual process but a relational one shaped by the social environment in which psychological change takes root.

Ultimately, this work suggests that mental health underpins the development of agency. Specifically, improvements in mental health can shift individuals' sense of self-efficacy, aspirations and choice sets, expanding the set of actions an individual views as feasible. Whether this renewed agency yields measurable empowerment depends on the surrounding environment of constraints, including one's outside options and prevailing social norms. Embedding these psychological mechanisms within economic models of empowerment advances our understanding of who benefits from interventions that aim to enhance economic empowerment, and clarifies the conditions under which psychological empowerment translates into sustained improvements in welfare and autonomy.

References

Gallis, J. A., Maselko, J., O'Donnell, K., Song, K., Saqib, K., Turner, E. L., & Sikander, S. (2018). Criterion-related validity and reliability of the urdu version of the patient health questionnaire in a sample of community-based pregnant women in pakistan. *PeerJ*, *2018*. https://doi.org/10.7717/peerj.5185

Gorman, L. L., O'Hara, M. W., Figueiredo, B., Hayes, S., Jacquemain, F., Kammerer, M. H., Klier, C. M., Rosi, S., Seneviratne, G., & Sutter-Dallay, A.-L. (2004). Adaptation of the structured clinical interview for dsm-iv disorders for assessing depression in women during pregnancy and post-partum across countries and cultures. *British Journal of Psychiatry*, *184*, s17–s23. https://doi.org/10.1192/bjp.184.46.s17

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The phq-9. *Journal of General Internal Medicine*, *16*, 606–613. https://doi.org/https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Maselko, J., Sikander, S., Turner, E. L., Bates, L. M., Ahmad, I., Atif, N., Baranov, V., Bhalotra, S., Bibi, A., Bibi, T., Bilal, S., Biroli, P., Chung, E., Gallis, J. A., Hagaman, A., Jamil, A., LeMasters, K., O'Donnell, K., Scherer, E., … Rahman, A. (2020). Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: A cluster randomised trial in pakistan. *The Lancet Psychiatry*, *7*, 775–787. https://doi.org/10.1016/S2215-0366(20)30258-3

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, *99*(2), 210–221. https://doi.org/10.1016/j.jdeveco.2012.01.002

NIPS & ICF. (2013). Pakistan demographic and health survey 2012-2013. https://dhsprogram.com/pubs/pdf/FR290/FR290.pdf

Sevim, D., Baranov, V., Bhalotra, S., Maselko, J., & Biroli, P. (2024). Trajectories of early childhood skill development and maternal mental health. *Journal of Human Resources*, *59*, S365–S401. https://doi.org/10.3368/jhr.1222-12693R3

Sikander, S., Ahmad, I., Atif, N., Zaidi, A., Vanobberghen, F., Weiss, H. A., Nisar, A., Tabana, H., Ain, Q. U., Bibi, A., Bilal, S., Bibi, T., Liaqat, R., Sharif, M., Zulfiqar, S., Fuhr, D. C., Price, L. N., Patel, V., & Rahman, A. (2019). Delivering the thinking healthy programme for perinatal depression through volunteer peers: A cluster randomised controlled trial in pakistan. *The Lancet Psychiatry*, *6*, 128–139. https://doi.org/10.1016/S2215-0366(18)30467-X

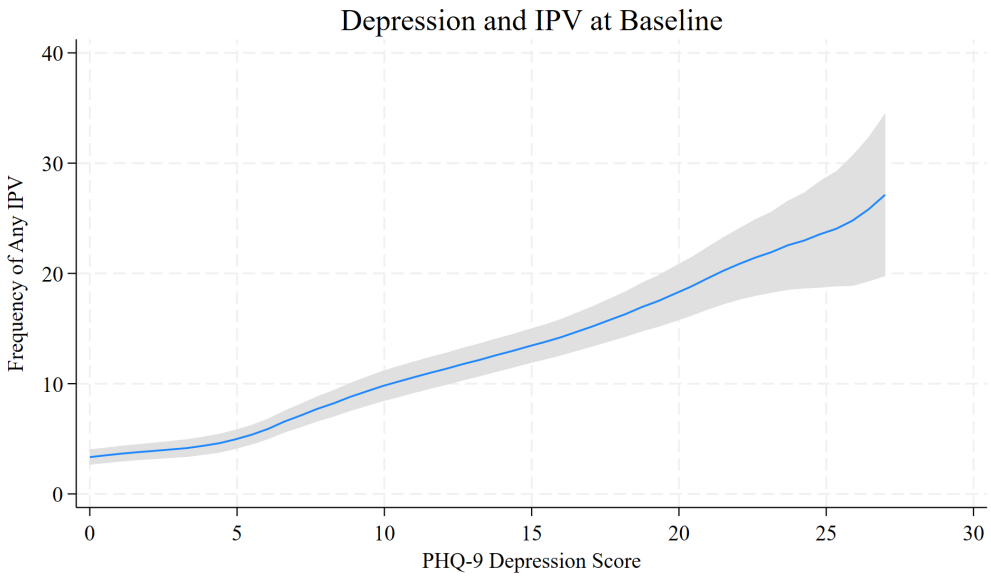
Sikander, S., Ahmad, I., Bates, L. M., Gallis, J., Hagaman, A., O'Donnell, K., Turner, E. L., Zaidi, A., Rahman, A., & Maselko, J. (2019). Cohort profile: Perinatal depression and child socioemotional development; the bachpan cohort study from rural pakistan. *BMJ Open*, *9*, 1–11. https://doi.org/10.1136/bmjopen-2018-025644

Turner, E. L., Sikander, S., Bangash, O., Zaidi, A., Bates, L., Gallis, J., Ganga, N., O'Donnell, K., Rahman, A., & Maselko, J. (2016). The effectiveness of the peer-delivered thinking healthy plus (thpp+) program for maternal depression and child socioemotional development in pakistan: Study protocol for a randomized controlled trial. *Trials*, *17*, 1–11. https://doi.org/10.1186/s13063-016-1530-y

WHO. (2005). Who multi-country study on women's health and domestic violence against women: Summary report. https://www.who.int/publications/i/item/9241593512

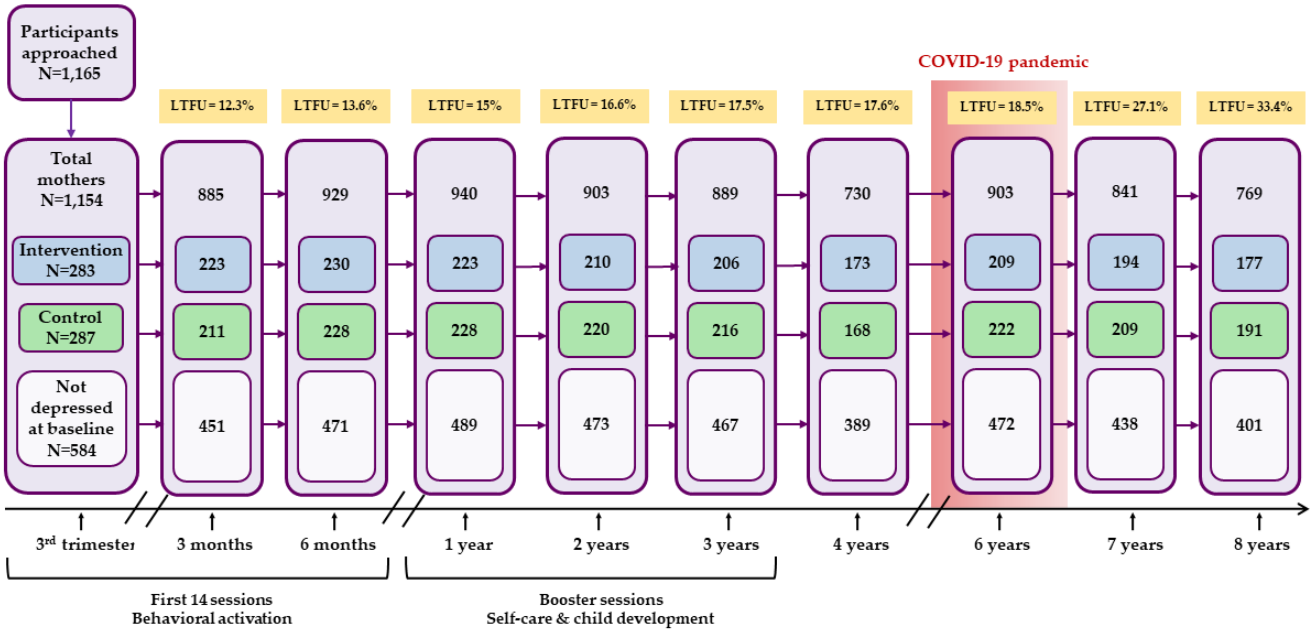
Figures

Figure 1: Relationship between depression and IPV at baseline



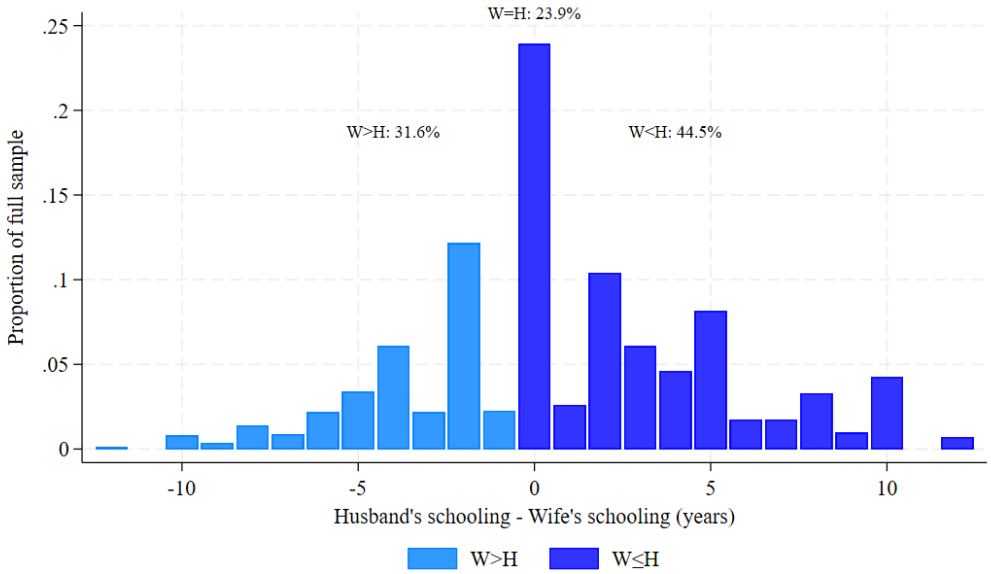
Notes: This figure plots a local polynomial regression of IPV frequency (of any type) against PHQ-9 depression scores at baseline. Shaded areas show 95% confidence intervals.

Figure 2: Timeline of THPP+ Intervention and Follow-ups



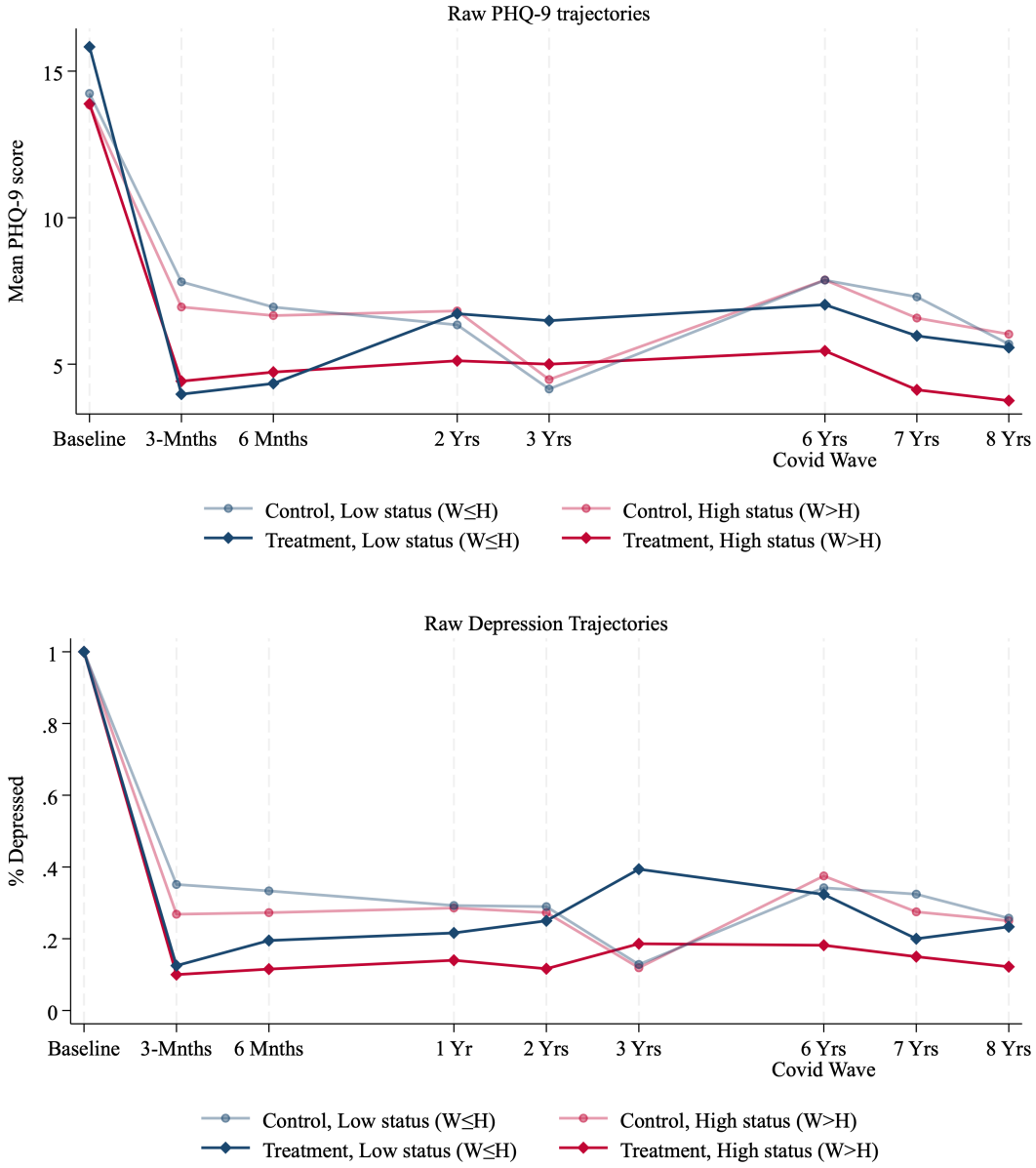
Notes: In year 4, surveys were collected using mobile phones, whilst all other survey waves took place through in-person interviews.

Figure 3: Distribution of baseline schooling gap between spouses



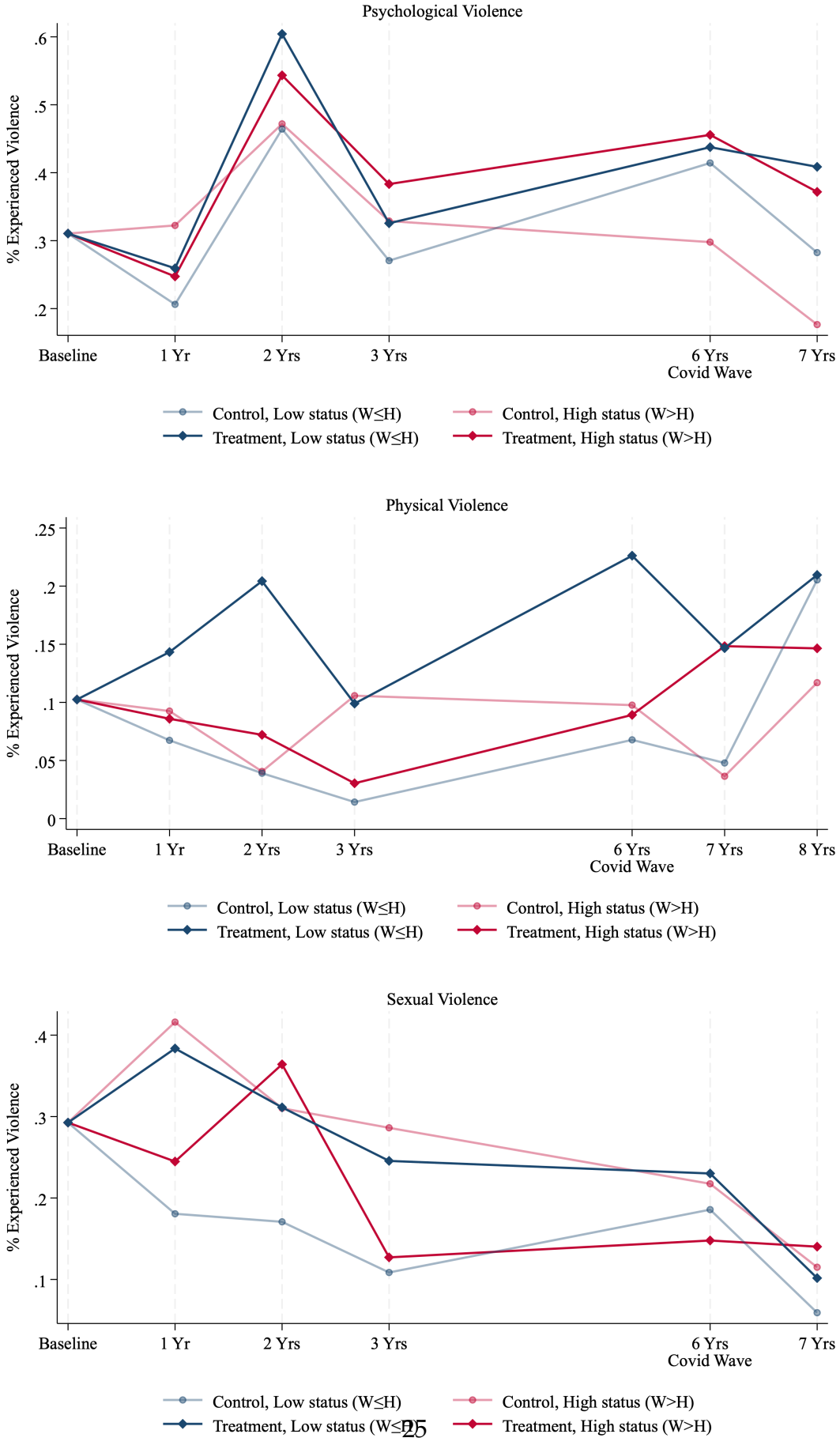
Notes: This figure plots the distribution of the gap in schooling years between spouses (Husband's schooling - Wife's schooling). A positive value indicates the husband has more schooling than the wife, the majority of the sample, and a negative value indicates the wife has more schooling than her husband.

Figure 4: Trajectories of depression by arm and spousal gap - high schooling women



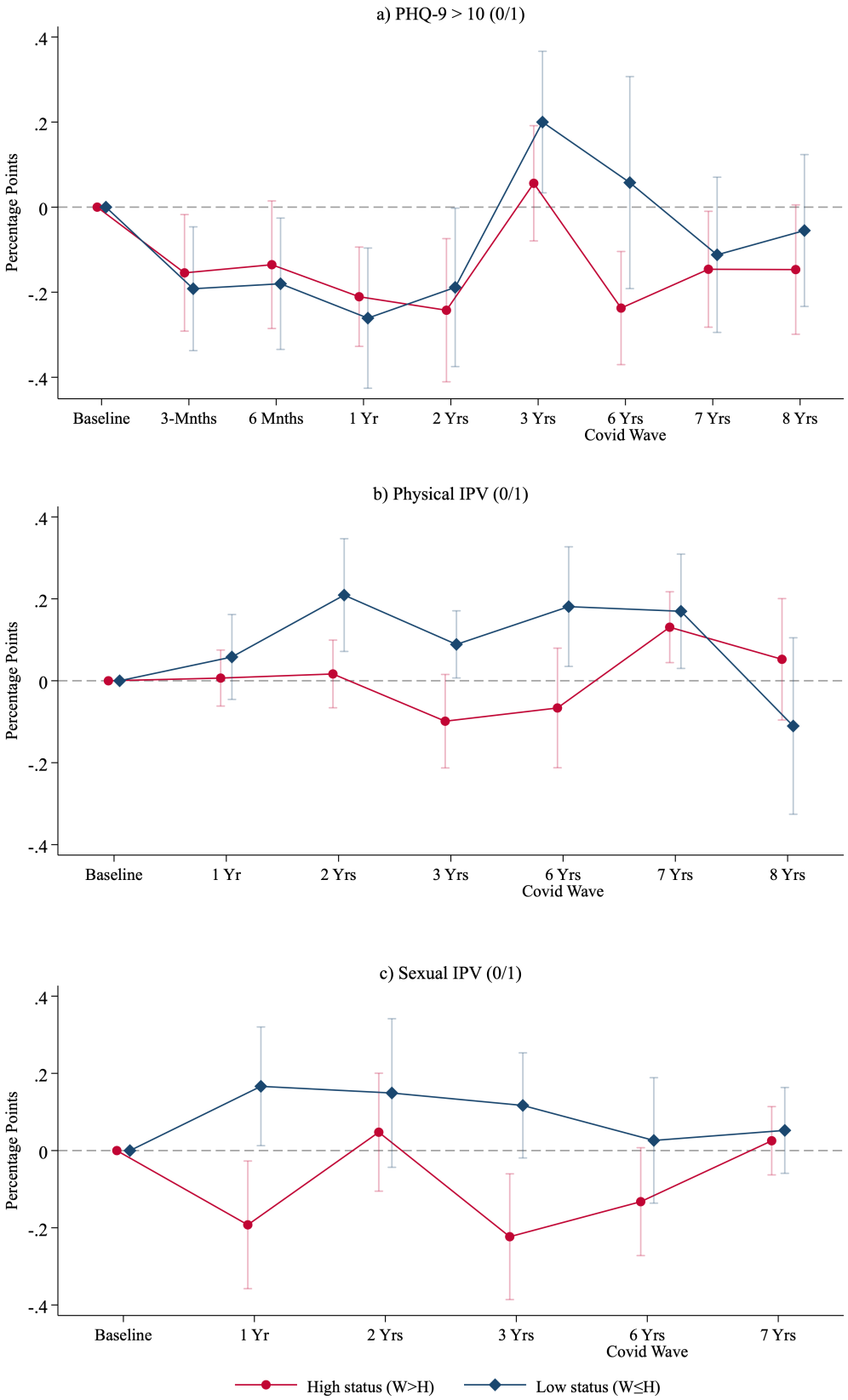
Notes: The above figures plot the raw trajectories of PHQ-9 scores and proportion of depressed women in the sample by treatment arm and women's status relative to their husbands over the eight years of follow-up, restricted to women with high school or more education.

Figure 5: Trajectories of IPV by arm and spousal gap - high schooling women



Notes: The above figures plot the raw trajectories of IPV: proportion of women experiencing psychological, physical and sexual in the sample by treatment arm and women's status relative to their husbands over the eight years of follow-up, restricted to women with high school or more education.

Figure 6: Trajectories of depression and IPV by spousal gap - high schooling women



Notes: Figures (a)-(d) plot the treatment effects on mental health, physical and sexual IPV by spousal schooling gap. Year 1 in figure (a) was imputed using SCID.

Tables

Table 1: Baseline balance across treatment arms

	Baseline Sample N=570				
	Control Mean (1)	SD (2)	Treatment Mean (3)	Diff T-C (4)	*p*-val (5)
Mother's age	27.29	(5.0)	26.80	-0.49	0.23
Mother's height (cm)	156.33	(6.1)	157.43	1.10	0.03
Mother's weight (kg)	61.24	(12.9)	60.17	-1.07	0.31
Mother's waist circ. (cm)	95.39	(10.4)	93.60	-1.79	0.05
Mother's blood pressure	71.82	(14.1)	70.66	-1.16	0.29
PHQ-9 Score	14.48	(3.6)	14.89	0.41	0.18
WHODAS Score	16.11	(9.1)	16.71	0.60	0.42
PSS Score	22.90	(7.5)	23.84	0.94	0.14
Current major dep. episode	0.73	(0.4)	0.78	0.05	0.21
Nuclear family	0.25	(0.4)	0.25	0.00	0.92
Joint/extended family	0.63	(0.5)	0.58	-0.05	0.18
Grandmother present	0.67	(0.5)	0.63	-0.04	0.36
Total adults in the household	5.70	(3.0)	5.33	-0.37	0.13
First child	0.25	(0.4)	0.23	-0.02	0.55
Number of boys	0.79	(1.0)	0.86	0.07	0.40
Number of girls	0.85	(1.1)	0.96	0.10	0.26
People per room	2.47	(1.9)	2.79	0.32	0.05
SES asset index	-0.32	(1.7)	-0.56	-0.24	0.09
Mother's education	6.80	(4.5)	6.83	0.03	0.95
Father's education	8.33	(3.3)	7.85	-0.48	0.10
W $>$ schooling than H	0.29	(0.5)	0.31	0.03	0.51
Any psychological IPV past 12m	0.49	(0.5)	0.52	0.03	0.47
Any physical IPV past 12m	0.23	(0.4)	0.27	0.03	0.39
Any sex-related IPV past 12m	0.37	(0.5)	0.43	0.07	0.10
Frequency of psychological IPV	13.54	(24.7)	14.46	0.92	0.68
Frequency of physical IPV	1.66	(5.1)	2.30	0.64	0.19
Frequency of sex-related IPV	5.12	(10.0)	7.01	1.90	0.06
Joint test (*p*-value)					0.09

Note: This table tests for balance in baseline characteristics. Columns 1 and 3 show the mean in the control, and treatment group at baseline.

Table 2: Treatment effects on mental health and violence

	Control group mean (1)	β (SE) (2)	Observations (3)
Panel A: Depression			
PHQ-9 Score	7.04	-1.117***	3548
		(0.243)	
PHQ-9 $>$ 10 (0/1)	0.30	-0.067***	3548
		(0.016)	
SCID (0/1)	0.20	-0.055***	3995
		(0.013)	
Panel B: Violence (0/1)			
Any Type	0.47	0.042*	2584
		(0.023)	
Psychological	0.41	0.041*	2581
		(0.021)	
Physical	0.13	0.035*	2815
		(0.019)	
Sexual	0.25	-0.014	2491
		(0.018)	

Note: Physical IPV (probability and frequency) include data from years 1-8. Estimates for psychological and sexual IPV only have data from years 1-7. Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p$<$0.01, ** p$<$0.05, * p$<$0.1.

Table 3: Mental health effects by women's absolute and relative schooling

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: High Schooling							
PHQ-9 Score	6.48	-2.085***	724	6.58	-1.798**	604	0.746
		(0.518)			(0.702)		
		[0.001]			[0.001]		
PHQ-9 $>$ 10 (0/1)	0.26	-0.145***	724	0.29	-0.086	604	0.395
		(0.036)			(0.058)		
		[0.001]			[0.023]		
SCID (0/1)	0.20	-0.131***	815	0.18	-0.093***	681	0.427
		(0.033)			(0.032)		
		[0.001]			[0.001]		
Panel B: Low Schooling							
PHQ-9 Score	10.12	-2.470	318	7.00	-0.027	1903	0.109
		(1.551)			(0.358)		
		[0.148]			[0.959]		
PHQ-9 $>$ 10 (0/1)	0.48	-0.142	318	0.30	0.005	1903	0.112
		(0.093)			(0.026)		
		[0.161]			[0.959]		
SCID (0/1)	0.38	-0.143**	360	0.18	0.006	2140	0.029
		(0.066)			(0.020)		
		[0.048]			[0.959]		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. N reports the effective sample size of each regression. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in column (2) and (5). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf p-values are reported in squared brackets under the SEs.

Table 4: IPV effects by spousal gap among high schooling women

Mi

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Psychological Violence	0.39	0.053	517	0.37	0.096*	444	0.589
		(0.048)			(0.052)		
		[0.297]			[0.075]		
Physical Violence	0.12	0.008	557	0.07	0.121**	481	0.055
		(0.038)			(0.045)		
		[0.763]			[0.020]		
Sexual Violence	0.26	-0.087**	499	0.16	0.111**	432	0.002
		(0.032)			(0.052)		
		[0.020]			[0.051]		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. N reports the effective sample size of each regression. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in column (2) and (5). Physical IPV (probability and frequency) include data from years 1-8. Estimates for psychological and sexual IPV only have data from years 1-7. Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf p-values are reported in squared brackets under the SEs.

Table 5: Agency effects by spousal gap among high schooling women

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Fertility							
Pregnant	0.10	-0.030	413	0.09	0.033	349	0.089
		(0.024)			(0.022)		
		[0.567]			[0.550]		
# Children Alive since index (Yr 8)	1.25	-0.354*	80	1.24	-0.171	67	0.566
		(0.201)			(0.220)		
		[0.392]			[0.776]		
Panel B: Decision-Making							
Decision-making z-score	-0.03	0.282***	348	0.03	-0.209*	296	0.001
		(0.089)			(0.117)		
		[0.007]			[0.137]		
Woman is involved in all	0.24	0.128***	348	0.26	-0.029	296	0.006
		(0.039)			(0.033)		
		[0.005]			[0.398]		
Panel C: Labor Force Participation							
Worked in last 12 months	0.10	0.062*	348	0.10	-0.069***	296	0.002
		(0.036)			(0.022)		
		[0.048]			[0.008]		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in columns (2) and (4). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf p-values are reported in squared brackets under the SEs.

Table 6: Alternative Mechanisms - High schooling women only

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Social Capital							
Interactions w.o. preg. wom.	1.58	0.125	187	1.47	0.002	157	0.486
		(0.174)			(0.125)		
Support from friends z-score	-0.03	-0.002	517	-0.13	0.145	433	0.342
		(0.106)			(0.118)		
Panel B: Gender Norms							
Gender norms z-score	0.15	0.021	81	0.14	0.308	65	0.361
		(0.199)			(0.187)		
Panel C: Economic Preferences							
Willing to take risk	4.40	0.196	81	5.11	-0.571	65	0.364
		(0.641)			(0.721)		
Willing to delay gratification	4.65	0.235	81	4.97	-0.037	65	0.766
		(0.633)			(0.796)		
Panel D: Parental Investment - Boys							
Infancy z-score	0.14	0.056	148	0.19	0.108	99	0.775
		(0.123)			(0.121)		
Early and Middle Childhood z-score	0.15	0.192	83	0.35	-0.050	56	0.461
		(0.256)			(0.242)		
Panel D: Parental Investment - Girls							
Infancy z-score	0.22	-0.378***	120	0.16	0.059	128	0.007
		(0.128)			(0.091)		
Early and Middle Childhood z-score	0.16	0.431	67	0.41	-0.183	71	0.136
		(0.368)			(0.183)		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in columns (2) and (4). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

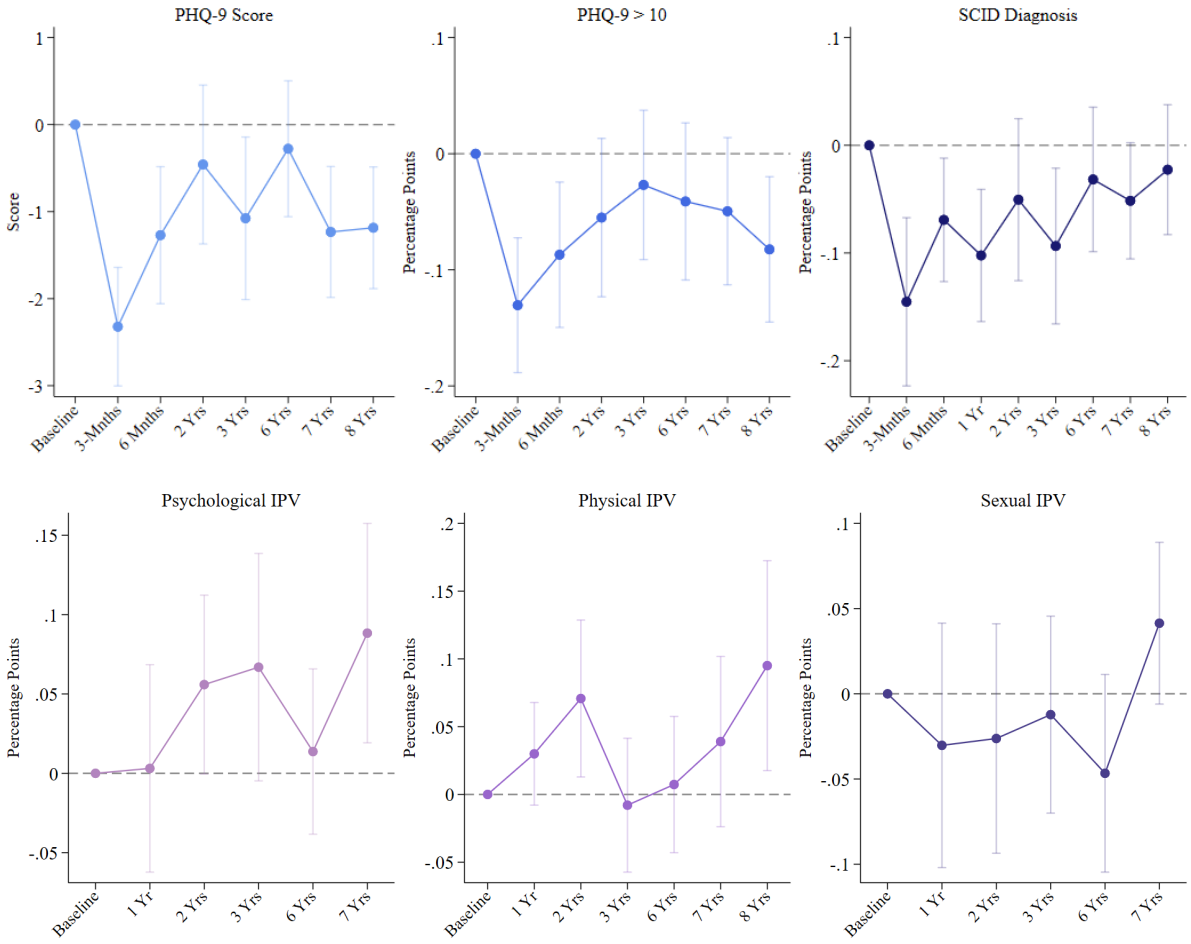
Table 7: Child Outcomes - High Schooling Women

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Child Outcomes - Boys							
Behavioral problems z-score	0.18	0.017	183	0.05	0.240	120	0.541
		(0.246)			(0.300)		
Literacy z-score	0.58	-0.112	87	0.43	0.075	59	0.576
		(0.228)			(0.282)		
Numeracy z-score	0.78	-0.567***	87	0.58	-0.262	59	0.354
		(0.204)			(0.238)		
Panel B: Child Outcomes - Girls							
Behavioral problems z-score	-0.26	-0.145	147	-0.01	0.187	156	0.175
		(0.197)			(0.151)		
Literacy z-score	0.74	0.946**	73	0.70	-0.147	73	0.107
		(0.434)			(0.428)		
Numeracy z-score	0.55	1.102***	73	0.69	-0.354	73	0.003
		(0.275)			(0.327)		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in columns (2) and (4). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

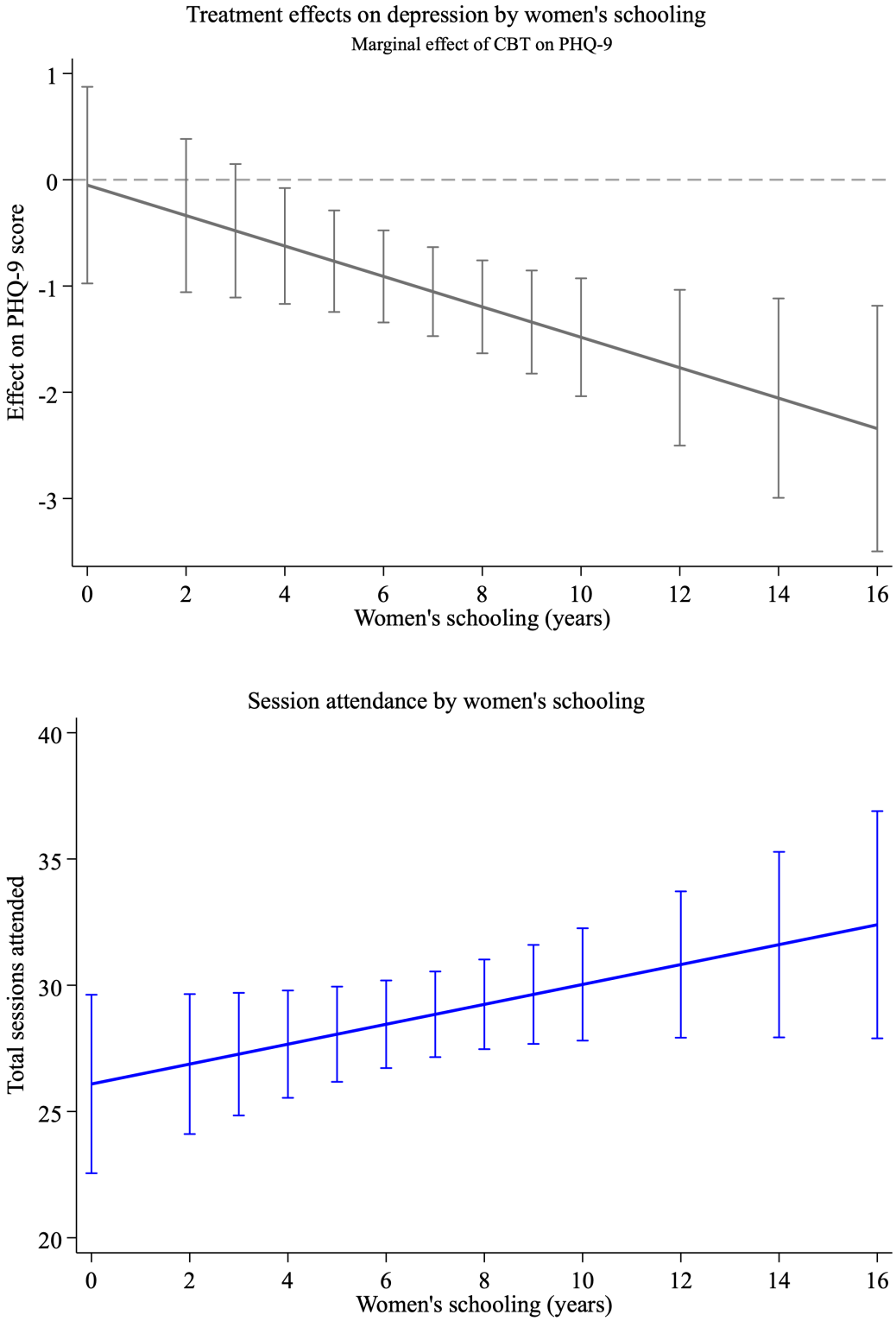
Appendix A - Additional Figures

Figure A1: Trajectories of Depression and IPV



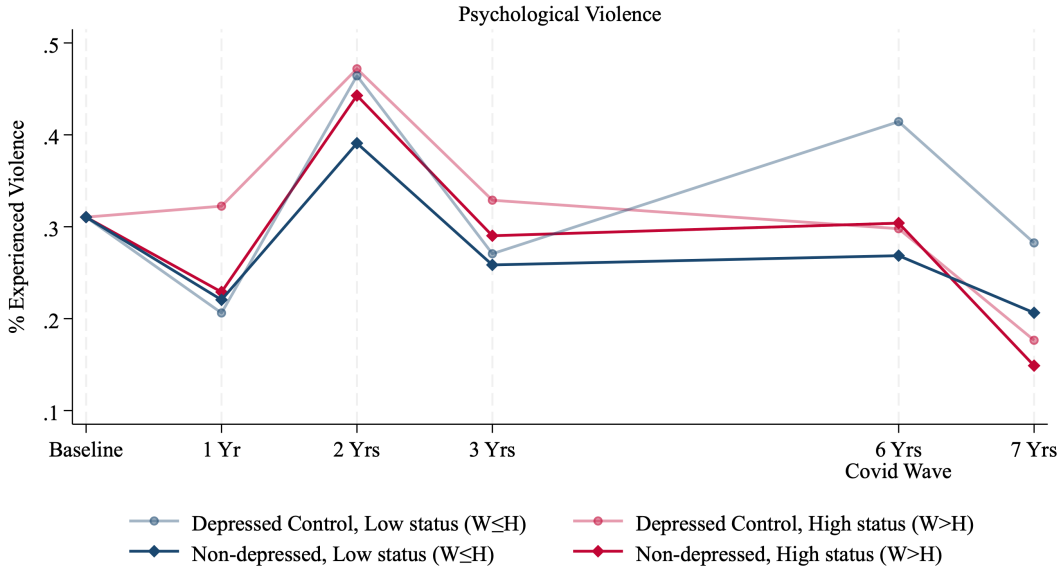
Notes: The above figures plot differences in trajectories between treatment and control arms for depression and the probability of experiencing IPV over time.

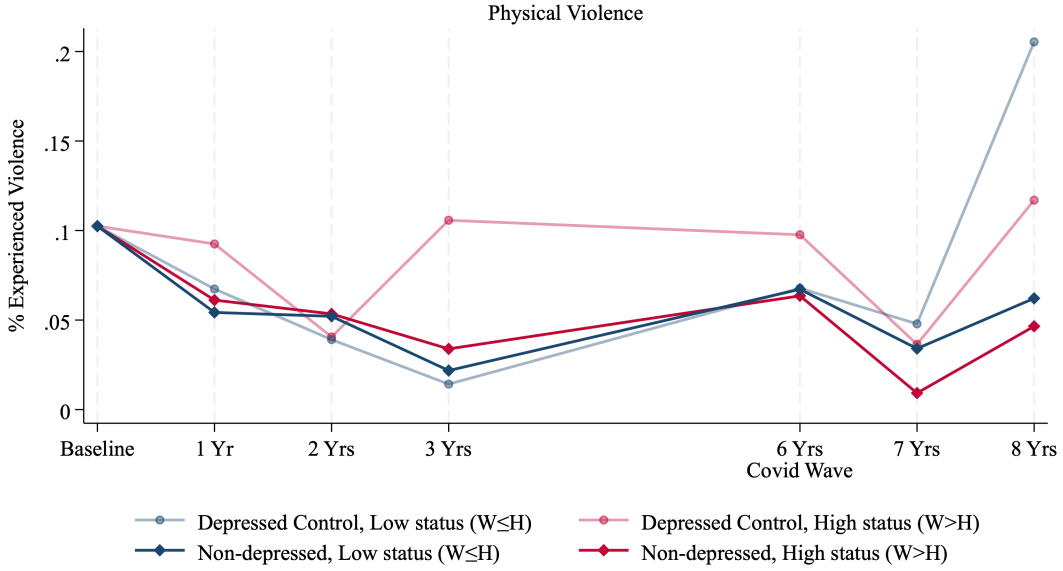
Figure A2: Treatment effects of depression by women's absolute schooling

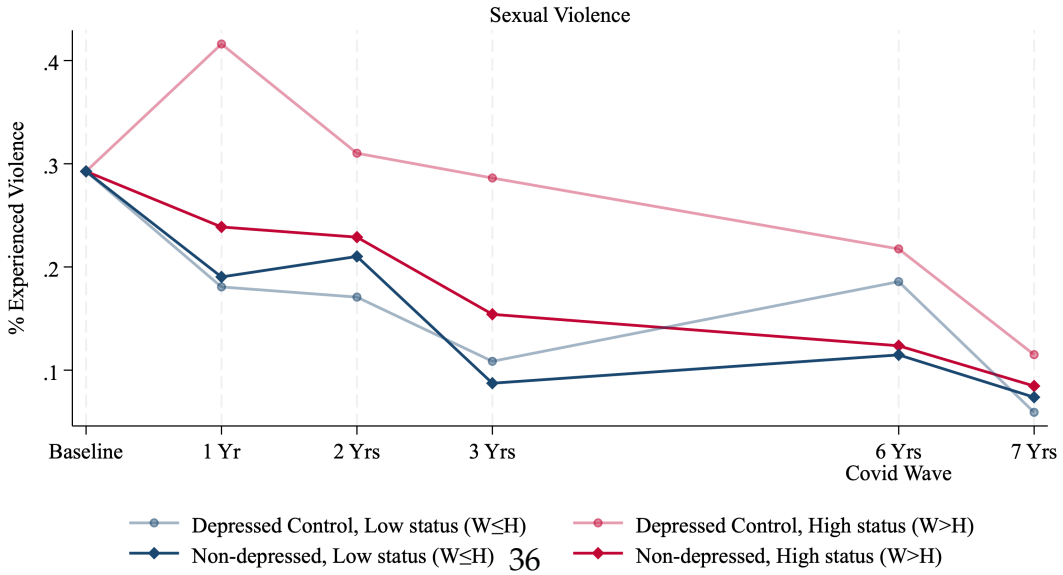


Notes: The top figure plots marginal treatment effects of CBT on PHQ-9 scores by women's absolute schooling levels with 90% CI. The bottom figure tests the correlation between the number of psychotherapy sessions attended and women's schooling levels.

Figure A3: Trajectories of IPV by spousal gap - non depressed and depressed control







Notes: The above figures plot the raw trajectories of IPV: proportion of women experiencing psychological, physical and sexual in the sample by women's status relative to their husbands over the eight years of follow-up, restricted to women with high school or more education and comparing the non-depressed cohort to the depressed control arm.

Appendix B – Additional Tables

B1 Main study measures used in the analyses 38
B2 Balance in baseline characteristics across follow-up waves 39
B3 Characteristics of attritors over time by treatment arm 40
B4 Treatment effects on mental health over time 41
B5 Treatment effects on violence over time 42
B6 Baseline characteristics by relative and absolute schooling 43
B7 Post baseline characteristics of women in the control arm 44
B8 Correlation between spousal schooling gap and female agency 45
B9 Fertility since index child by schooling gap 46
B10 Baseline balance across relative and absolute schooling groups 47
B11 Violence effects by women's absolute and relative schooling 48
B12 WHO IPV Items by schooling gap - high schooling women only 49
B13 Effects on relationship quality - high schooling women only 50
B14 Characteristics of Peer Volunteers 51
B15 Do patient characteristics predict PV characteristics? 52
B16 Treatment effects for W>H by PV characteristics 53
B17 Treatment effects for W≤H by PV characteristics 54

Appendix Table B1: Main study measures used in the analyses

Variable	Description
Panel A: Mental Health	
PHQ-9 Score	The Patient Health Questionnaire (PHQ-9) is a 12-item module measured in every survey wave, except year 1 of follow-up. It includes questions on frequency of symptoms such as feeling tired, having poor appetite, trouble sleeping, etc., with a reference of period of 2 weeks preceding the survey.
Panel B: Intimate Partner Violence	
Psychological IPV	Has this happened to you in the past 12 months? If yes, with what frequency? (1-2 times, 3-5, 6-10, 11-20, >20) -Has your husband ever insulted you or made you feel bad about yourself? -Done things that scared or intimidated you on purpose, for example, by the way he looked at you, or by yelling or smashing things? -Threatened to hurt you or someone you care about? -Called you ugly or said something else negative about your appearance? -Destroyed something belonging to you on purpose? -Threatened to take another wife? -Threatened to abandon you or send you back to your natal family? -Threatened to divorce you? -Said you were not able to please him sexually. **Any:** Binary variable = 1 if respondent experienced any of the 9 acts.
Physical IPV	Has this happened to you in the past 12 months? If yes, with what frequency? (1-2 times, 3-5, 6-10, 11-20, >20) -Slapped you or thrown something at you that could hurt you? Pushed you or shoved you or pulled your hair? -Choked or burnt you on purpose? -Threatened to use a gun, knife or other weapon against you? -Actually used a gun, knife or other weapon against you? **Any:** Binary variable = 1 if respondent experienced any of the 4 acts.
Sexual IPV	Has this happened to you in the past 12 months? If yes, with what frequency? (1-2 times, 3-5, 6-10, 11-20, >20) -Did your husband ever physically force you to have sexual intercourse when you did not want to? -Did you ever have sexual intercourse when you did not want to because you were afraid of what your husband might do? -Did your husband ever force you to do something sexual that you found degrading or humiliating? **Any:** Binary variable = 1 if respondent experienced any of the 3 acts.

Note: Measures of intimate partner violence were collected at baseline, years 1 – 3 and years 6 – 7.

Appendix Table B2: Balance in baseline characteristics across follow-up waves

	6 Mths N=940		Year 1 N=940		Year 2 N=903		Year 3 N=889		Year 6 N=898		Year 7 N=841		Year 8 N=841	
	Diff T-C (1)	*p*-val (2)	Diff T-C (3)	*p*-val (4)	Diff T-C (5)	*p*-val (6)	Diff T-C (7)	*p*-val (8)	Diff T-C (9)	*p*-val (10)	Diff T-C (11)	*p*-val (12)	Diff T-C (13)	*p*-val (14)
Mother's age	-0.59	0.18	-0.36	0.42	-0.31	0.49	-0.31	0.48	-0.49	0.28	-0.47	0.31	-0.60	0.21
Mother's height (cm)	1.00	0.08	1.04	0.07	1.35	0.02	1.02	0.09	1.17	0.05	1.48	0.02	1.47	0.02
Mother's weight (kg)	-1.21	0.32	-0.86	0.48	-1.06	0.40	-1.29	0.31	-1.43	0.25	-1.30	0.31	-1.37	0.29
Mother's waist circ. (cm)	-2.02	0.04	-1.59	0.11	-2.14	0.04	-1.75	0.09	-2.10	0.04	-2.34	0.03	-2.34	0.03
Mother's blood pressure	-0.03	0.98	-0.85	0.48	-0.03	0.98	-0.73	0.55	-0.61	0.62	-0.76	0.54	-0.68	0.59
PHQ-9 Score	0.40	0.24	0.48	0.16	0.36	0.30	0.38	0.29	0.27	0.44	0.17	0.65	-0.01	0.98
WHODAS Score	0.60	0.46	0.81	0.33	1.00	0.24	0.81	0.34	0.63	0.46	0.44	0.62	0.31	0.72
PSS Score	1.31	0.07	1.11	0.12	1.22	0.10	1.02	0.16	1.05	0.15	1.09	0.15	0.92	0.23
Current major dep. episode	0.05	0.18	0.04	0.35	0.04	0.31	0.04	0.30	0.05	0.28	0.03	0.43	0.03	0.55
Nuclear family	0.01	0.77	0.01	0.90	0.01	0.88	0.02	0.62	-0.02	0.62	-0.02	0.68	-0.02	0.65
Joint/extended family	-0.06	0.20	-0.05	0.29	-0.06	0.20	-0.07	0.11	-0.04	0.40	-0.05	0.32	-0.05	0.34
Grandmother present	-0.03	0.53	-0.03	0.51	-0.04	0.34	-0.03	0.45	-0.01	0.84	-0.02	0.60	-0.01	0.78
Total adults in the household	-0.32	0.23	-0.33	0.23	-0.21	0.44	-0.28	0.31	-0.08	0.77	-0.24	0.40	-0.25	0.38
First child	-0.00	0.96	0.01	0.90	-0.01	0.76	-0.02	0.58	-0.00	0.97	-0.00	0.93	0.00	1.00
Number of boys	0.00	0.98	0.02	0.80	0.04	0.69	0.04	0.68	-0.05	0.57	-0.04	0.64	-0.09	0.31
Number of girls	0.07	0.49	0.09	0.37	0.12	0.23	0.05	0.62	0.05	0.63	0.06	0.60	0.05	0.62
People per room	0.33	0.04	0.35	0.03	0.33	0.04	0.32	0.06	0.35	0.04	0.36	0.04	0.30	0.07
SES asset index	-0.13	0.40	-0.13	0.42	-0.16	0.33	-0.17	0.30	-0.08	0.61	-0.13	0.45	-0.01	0.95
Mother's education	0.38	0.35	0.31	0.45	0.09	0.84	0.10	0.82	0.24	0.56	-0.02	0.95	0.09	0.84
Father's education	-0.56	0.08	-0.61	0.06	-0.84	0.01	-0.75	0.02	-0.61	0.06	-0.82	0.02	-0.82	0.02
W > schooling than H	0.05	0.18	0.05	0.19	0.04	0.29	0.04	0.34	0.04	0.32	0.04	0.32	0.05	0.27
Any psychological IPV past 12m	0.02	0.60	0.00	0.95	0.02	0.65	0.01	0.77	-0.00	0.94	-0.02	0.75	-0.01	0.85
Any physical IPV past 12m	0.03	0.43	0.02	0.60	0.05	0.23	0.05	0.29	0.05	0.28	0.04	0.40	0.04	0.41
Any sex-related IPV past 12m	0.09	0.04	0.08	0.07	0.11	0.02	0.09	0.06	0.10	0.04	0.09	0.08	0.11	0.03
Frequency of psychological IPV	0.15	0.95	0.09	0.97	0.19	0.94	0.38	0.89	-0.00	1.00	-0.25	0.93	-0.18	0.95
Frequency of physical IPV	0.93	0.11	1.00	0.07	0.94	0.10	1.13	0.05	1.03	0.08	1.07	0.06	1.32	0.01
Frequency of sex-related IPV	2.72	0.02	2.55	0.02	3.00	0.01	3.06	0.01	2.90	0.01	3.20	0.01	3.72	0.00
Joint test (*p*-value)		0.17		0.17		0.09		0.22		0.22		0.03		0.00

Note: This table tests for balance in baseline characteristics between arms across follow-up waves.

Appendix Table B3: Characteristics of attritors over time by treatment arm

	6 Months		Year 1		Year 2		Year 3		Year 6		Year 7		Year 8	
	Diff T-C (1)	*p*-val (2)	Diff T-C (3)	*p*-val (4)	Diff T-C (5)	*p*-val (6)	Diff T-C (7)	*p*-val (8)	Diff T-C (9)	*p*-val (10)	Diff T-C (11)	*p*-val (12)	Diff T-C (13)	*p*-val (14)
Mother's age	0.06	0.95	0.95	0.31	1.04	0.25	0.97	0.27	0.41	0.63	0.48	0.54	0.14	0.85
Mother's height (cm)	-1.49	0.17	-1.34	0.22	-0.37	0.70	-1.36	0.16	-0.94	0.35	-0.23	0.80	-0.29	0.75
Mother's weight (kg)	0.71	0.74	1.79	0.41	0.79	0.68	0.22	0.91	-0.22	0.92	0.33	0.86	0.15	0.93
Mother's waist circ. (cm)	1.08	0.59	2.46	0.23	0.44	0.80	1.72	0.32	0.71	0.71	0.36	0.83	0.40	0.81
Mother's blood pressure	5.62	0.04	2.34	0.38	4.73	0.06	2.45	0.31	3.03	0.22	2.34	0.29	2.50	0.24
PHQ-9 Score	-0.47	0.49	-0.12	0.86	-0.48	0.43	-0.47	0.42	-0.79	0.20	-0.94	0.10	-1.29	0.03
WHODAS Score	-0.67	0.70	0.22	0.90	0.63	0.68	-0.03	0.98	-0.47	0.76	-0.89	0.52	-1.12	0.42
PSS Score	0.57	0.68	-0.30	0.82	-0.09	0.95	-0.78	0.54	-0.62	0.63	-0.55	0.63	-0.89	0.44
Current major dep. episode	-0.02	0.82	-0.07	0.33	-0.04	0.50	-0.04	0.51	-0.04	0.58	-0.06	0.32	-0.08	0.18
Nuclear family	0.01	0.88	0.01	0.95	0.01	0.89	0.05	0.50	-0.06	0.43	-0.05	0.52	-0.05	0.51
Joint/extended family	0.05	0.59	0.08	0.40	0.03	0.69	-0.01	0.92	0.09	0.27	0.06	0.40	0.07	0.37
Grandmother present	0.09	0.36	0.06	0.52	0.01	0.95	0.03	0.69	0.11	0.20	0.05	0.49	0.08	0.28
Total adults in the household	0.61	0.29	0.53	0.36	0.84	0.12	0.60	0.25	1.24	0.02	0.67	0.17	0.64	0.18
First child	0.10	0.26	0.12	0.14	0.05	0.50	0.02	0.76	0.09	0.24	0.07	0.29	0.08	0.24
Number of boys	-0.34	0.09	-0.24	0.22	-0.15	0.37	-0.15	0.36	-0.43	0.01	-0.33	0.04	-0.44	0.01
Number of girls	-0.24	0.24	-0.15	0.48	-0.03	0.89	-0.24	0.20	-0.26	0.21	-0.21	0.25	-0.22	0.23
People per room	-0.33	0.53	-0.20	0.68	-0.24	0.58	-0.30	0.46	-0.21	0.62	-0.18	0.63	-0.31	0.41
SES asset index	0.72	0.03	0.65	0.04	0.47	0.10	0.41	0.16	0.69	0.02	0.48	0.08	0.74	0.01
Mother's education	1.47	0.11	1.05	0.23	0.18	0.82	0.17	0.83	0.72	0.38	-0.08	0.92	0.15	0.84
Father's education	0.20	0.77	-0.00	1.00	-0.70	0.28	-0.31	0.60	0.03	0.96	-0.38	0.51	-0.34	0.55
Any psychological IPV past 12m	-0.06	0.55	-0.14	0.14	-0.06	0.49	-0.09	0.29	-0.14	0.10	-0.15	0.06	-0.12	0.11
Any physical IPV past 12m	-0.03	0.73	-0.07	0.41	0.03	0.70	0.00	0.94	0.01	0.88	-0.02	0.77	-0.02	0.76
Any sex-related IPV past 12m	0.04	0.70	-0.01	0.91	0.05	0.54	-0.00	0.95	0.03	0.76	-0.02	0.75	0.02	0.76
Frequency of psychological IPV in the past 12 months	-0.21	0.37	-0.22	0.39	-0.18	0.45	-0.16	0.46	-0.22	0.33	-0.22	0.26	-0.20	0.35
Frequency of physical IPV in the past 12 months	0.07	0.51	0.09	0.53	0.04	0.76	0.09	0.42	0.07	0.56	0.06	0.65	0.13	0.33
Frequency of sexual IPV in the past 12 months	0.24	0.43	0.08	0.80	0.22	0.46	0.20	0.44	0.15	0.60	0.16	0.56	0.37	0.17
Joint test (*p*-value)		0.43		0.80		0.45		0.44		0.60		0.55		0.01

Note: This table tests for differential attrition across follow-up waves by treatment arm.

Appendix Table B4: Treatment effects on mental health over time

	PHQ-9 Score (1)	PHQ-9>10 (2)	SCID (3)
At the 3-month follow-up	-2.321***	-0.130***	-0.145**
	(0.404)	(0.034)	(0.046)
At the 6-month follow-up	-1.270**	-0.087*	-0.069*
	(0.467)	(0.037)	(0.034)
At the 1-year follow-up			-0.102**
			(0.036)
At the 2-year follow-up	-0.458	-0.055	-0.050
	(0.542)	(0.040)	(0.045)
At the 3-year follow-up	-1.076	-0.027	-0.094*
	(0.554)	(0.038)	(0.043)
At the 6-year follow-up	-0.277	-0.041	-0.032
	(0.463)	(0.040)	(0.040)
At the 7-year follow-up	-1.233**	-0.050	-0.052
	(0.446)	(0.038)	(0.032)
At the 8-year follow-up	-1.185**	-0.082*	-0.023
	(0.415)	(0.037)	(0.036)
Intervention × Post	-1.117***	-0.067***	-0.055***
	(0.243)	(0.016)	(0.013)
Observations	3548	3548	3995
Mean	6.775	0.283	0.196
Imbalanced Controls	Y	Y	Y
Controls × Wave FEs	Y	Y	Y
Wave FEs	Y	Y	Y
Union Council FEs	Y	Y	Y

Note: Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table B5: Treatment effects on violence over time

	Probability of IPV				Frequency of IPV			
	Any (1)	Psychological (2)	Physical (3)	Sexual (4)	Any (5)	Psychological (6)	Physical (7)	Sexual (8)
At the 1-year follow-up	0.020	0.003	0.030	-0.030	-0.673	0.187	0.088	-0.380
	(0.041)	(0.039)	(0.022)	(0.043)	(1.384)	(0.911)	(0.098)	(0.511)
At the 2-year follow-up	0.038	0.056	0.071*	-0.026	1.470	1.704	0.226	-0.294
	(0.038)	(0.033)	(0.034)	(0.040)	(1.030)	(0.926)	(0.129)	(0.487)
At the 3-year follow-up	0.085	0.067	-0.008	-0.012	-0.765	0.101	-0.067	0.061
	(0.049)	(0.043)	(0.029)	(0.034)	(0.730)	(0.745)	(0.096)	(0.344)
At the 6-year follow-up	-0.016	0.014	0.007	-0.047	-0.483	0.798	-0.038	-0.644*
	(0.030)	(0.031)	(0.030)	(0.034)	(1.004)	(0.965)	(0.088)	(0.305)
At the 7-year follow-up	0.081	0.088*	0.039	0.041	1.313	1.661	0.071	0.257
	(0.043)	(0.041)	(0.037)	(0.028)	(1.112)	(0.930)	(0.117)	(0.298)
At the 8-year follow-up			0.095*				0.108	
			(0.046)				(0.135)	
Intervention × Post	0.042*	0.041*	0.035*	-0.014	0.242	0.876	0.064	-0.210
	(0.023)	(0.021)	(0.019)	(0.018)	(0.605)	(0.618)	(0.077)	(0.253)
Observations	2584	2581	2815	2491	2476	2577	2812	2483
Mean	0.470	0.412	0.126	0.249	8.263	5.782	0.396	2.072
Imbalanced Controls	Y	Y	Y	Y	Y	Y	Y	Y
Controls × Wave FEs	Y	Y	Y	Y	Y	Y	Y	Y
Wave FEs	Y	Y	Y	Y	Y	Y	Y	Y
Union Council FEs	Y	Y	Y	Y	Y	Y	Y	Y

Note: In year 8 of follow-up, women were asked one single question on incidence of IPV: How often has this type of behavior (hitting, slapping, pushing, etc) happened to you? Responses were then recoded to match format used in years 1-7. Robust standard errors clustered at the unit of randomization are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix Table B6: Baseline characteristics by relative and absolute schooling

	High Schooling		Low Schooling	
	W > H (1)	W ≤ H (2)	W > H (3)	W ≤ H (4)
Mother's age	26.59	26.71	26.12	27.49
Mother's age at marriage	22.37	22.47	20.48	21.66
# of pregnancies	2.66	3.00	3.54	3.74
# of children	1.15	1.58	1.74	2.00
Index is first child	0.38	0.24	0.22	0.19
Woman's schooling years	12.20	10.23	5.86	3.75
Husband's schooling years	8.56	11.06	1.92	7.97
Woman is employed	0.07	0.06	0.04	0.07
Husband is employed	0.96	0.89	0.90	0.86
Poorest 50%	0.38	0.36	0.82	0.75
# people in hhd	8.18	8.23	8.38	8.83
PHQ-9 Score	13.87	15.01	15.16	14.83
SCID (0/1)	0.73	0.76	0.72	0.77
Decisions about husb salary	0.30	0.23	0.26	0.29
Decisions about own health	0.49	0.34	0.40	0.41
Decisions about major purchases	0.28	0.26	0.22	0.24
Decisions about minor purchases	0.27	0.22	0.22	0.29
Decisions about visits	0.19	0.23	0.28	0.25
Allowed to go to market	0.83	0.65	0.78	0.76
Allowed to go places outside village	0.79	0.73	0.70	0.74
Life satisfaction	2.65	2.59	2.84	2.85
MSPSS Family Score	4.58	4.33	3.70	4.08
MSPSS Friends Score	3.65	3.04	3.04	3.35
MSPSS Significant Other	4.80	4.54	3.66	4.08
Family < 1 hour away	0.64	0.70	0.59	0.57
Visits family 6+ times/year	0.72	0.73	0.57	0.59
Any psychological IPV past 12m	0.45	0.51	0.52	0.52
Any physical IPV past 12m	0.21	0.18	0.36	0.27
Any sex-related IPV past 12m	0.34	0.40	0.55	0.40
Any IPV in past 12 months	0.60	0.60	0.76	0.61
Frequency of psychological IPV	10.02	16.76	12.73	14.95
Frequency of physical IPV	1.15	2.06	1.86	2.30
Frequency of sex-related IPV	4.94	5.61	8.57	6.23
Frequency of any IPV in past 12m	16.19	24.71	23.46	23.81
Marital unrest	0.29	0.31	0.48	0.36

Note: The above table shows descriptive statistics at baseline by women's relative and absolute schooling, pooling treatment and control.

Appendix Table B7: Post baseline characteristics of women in the control arm

	High Schooling		Low Schooling	
	W > H	W ≤ H	W > H	W ≤ H
	(1)	(2)	(3)	(4)
PHQ-9 Score	6.20	6.30	9.45	6.78
SCID (0/1)	0.20	0.18	0.38	0.18
Decisions about husb salary	0.37	0.38	0.45	0.35
Decisions about own health	0.51	0.47	0.54	0.47
Decisions about major purchases	0.36	0.39	0.37	0.36
Decisions about minor purchases	0.38	0.41	0.46	0.40
Decisions about visits	0.43	0.41	0.46	0.41
Allowed to go to market	0.95	0.96	0.83	0.87
Allowed to go places outside village	0.88	0.95	0.85	0.80
# of girls alive	1.31	1.84	1.44	1.69
# of boys alive	1.30	1.32	1.62	1.63
MSPSS Family Score	5.45	5.34	4.70	5.39
MSPSS Friends Score	4.27	4.08	3.74	4.31
MSPSS Significant Other	5.53	5.62	4.66	5.39
Any psychological IPV past 12m	0.39	0.37	0.45	0.43
Any physical IPV past 12m	0.11	0.06	0.27	0.14
Any sex-related IPV past 12m	0.26	0.16	0.36	0.26
Any IPV in past 12 months	0.46	0.40	0.57	0.50
Frequency of psychological IPV	6.27	5.90	15.45	8.88
Frequency of physical IPV	1.16	1.06	4.27	1.22
Frequency of sex-related IPV	1.99	1.47	4.57	3.47
Frequency of any IPV in past 12m	9.25	8.66	24.84	13.37
Marital unrest	0.16	0.12	0.25	0.17

Note: The above table shows descriptive statistics post baseline by women's relative and absolute schooling, among the control arm only.

Appendix Table B8: Correlation between spousal schooling gap and female agency

	Decisions making index (1)	Decisions about own healthcare (2)	Decisions about major purchases (3)
Panel A: DHS - Rural Pakistan			
W > schooling than H	0.206**	0.084*	0.103**
	(0.079)	(0.040)	(0.038)
Observations	1637	1637	1637
Panel B: Bachpan			
W > schooling than H	0.193***	0.109***	0.055*
	(0.056)	(0.031)	(0.027)
Observations	1080	1154	1154
Panel C: Bachpan			
W > schooling than H	0.203**	0.123***	0.040
	(0.066)	(0.037)	(0.031)
Woman has sec. schooling+	0.046	0.001	0.058
	(0.068)	(0.037)	(0.032)
Husband has sec. schooling+	0.102	0.053	0.029
	(0.064)	(0.035)	(0.030)
Observations	1080	1154	1154
Panel D: Bachpan			
W > schooling than H	0.199***	0.112***	0.058*
	(0.058)	(0.032)	(0.027)
Moderate to severe depression	-0.117	-0.043	-0.032
	(0.123)	(0.068)	(0.058)
W > H × Depressed	-0.182	-0.084	-0.085
	(0.262)	(0.143)	(0.123)
Observations	1080	1154	1154

Note: In the DHS data, the standardized index in column one includes two variables: the woman's involvement in decisions about her own healthcare and major household purchases. In the Bachpan data, it also includes whether the woman has input in deciding how her husband's earnings are spent and whether she is allowed to go to the market. Robust standard errors at the cluster level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Appendix Table B9: Fertility since index child by schooling gap

	Additional number of children since birth of index child by Yr 8	
	W>H	W≤H
	(1)	(2)
Husband wants more children than wife (0/1)	0.080	0.250**
	(0.131)	(0.092)
Mother's desired fertility	0.212*	0.193**
	(0.092)	(0.056)
SES asset index	0.091	-0.112*
	(0.085)	(0.050)
Mother's age	-0.079***	-0.069***
	(0.018)	(0.010)
Mother's schooling years	-0.000	0.011
	(0.028)	(0.016)
Nuclear family	0.010	-0.145
	(0.140)	(0.097)
Outcome Mean	1.18	1.00
R-squared	0.291	0.239
Observations	256	558

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. Models control for Union Council fixed effects, wave fixed effects, and are restricted to the depressed control and non-depressed groups, which are also included as controls. Robust standard errors clustered at the unit of randomization are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Appendix Table B10: Baseline balance across relative and absolute schooling groups

	High Schooling				Low Schooling			
	W > H		W ≤ H		W > H		W ≤ H	
	Diff T-C (1)	*p*-val (2)	Diff T-C (3)	*p*-val (4)	Diff T-C (5)	*p*-val (6)	Diff T-C (7)	*p*-val (8)
Mother's age	-0.87	0.34	-0.22	0.78	1.43	0.33	-0.69	0.22
Mother's age at marriage	-0.29	0.17	-0.77	0.00	1.32	0.00	-0.07	0.65
# of pregnancies	0.09	0.74	0.08	0.80	0.68	0.20	-0.01	0.96
# of children	0.13	0.05	-0.01	0.95	0.65	0.00	0.17	0.00
Index is first child	-0.01	0.63	0.08	0.00	-0.24	0.00	-0.02	0.07
Woman's schooling years	0.13	0.88	-0.52	0.01	-0.12	0.79	0.30	0.28
Husband's schooling years	-0.94	0.03	-0.04	0.88	-0.33	0.57	-0.35	0.18
Woman is employed	0.01	0.80	0.05	0.03	-0.02	0.38	0.00	0.93
Husband is employed	-0.01	0.75	-0.02	0.54	-0.04	0.32	-0.02	0.44
Poorest 50%	-0.03	0.27	0.04	0.14	0.23	0.00	0.05	0.00
# people in hhd	-0.41	0.60	1.09	0.13	-1.79	0.24	-0.38	0.50
PHQ-9 Score	-1.44	0.00	-0.39	0.42	-2.44	0.00	0.46	0.11
PHQ-9 $>$ 10	-0.07	0.01	-0.03	0.30	-0.18	0.00	0.03	0.19
Decisions about husb salary	0.09	0.07	-0.05	0.35	0.09	0.26	0.01	0.84
Decisions about own health	0.01	0.78	-0.05	0.41	0.05	0.50	0.03	0.31
Decisions about major purchases	0.10	0.05	-0.06	0.24	0.18	0.02	0.02	0.42
Decisions about minor purchases	0.09	0.07	-0.09	0.09	0.11	0.15	-0.01	0.76
Decisions about visits	0.02	0.65	-0.05	0.35	0.03	0.72	0.00	0.89
Allowed to go to market	-0.05	0.15	-0.07	0.12	0.12	0.04	0.05	0.04
Allowed to go places outside village	-0.01	0.85	-0.09	0.03	0.02	0.81	0.04	0.09
Life satisfaction	0.55	0.55	-4.12	0.12	-3.56	0.43	-0.76	0.47
MSPSS Family Score	0.13	0.25	0.05	0.73	0.43	0.05	-0.12	0.13
MSPSS Friends Score	0.29	0.05	0.33	0.05	0.01	0.96	-0.11	0.26
MSPSS Significant Other	0.27	0.02	0.02	0.86	0.44	0.04	-0.03	0.72
Family $<$ 1 hour away	-0.06	0.08	-0.07	0.05	0.50	0.00	-0.10	0.00
Visits family 6+ times/year	0.01	0.65	-0.06	0.05	0.34	0.00	-0.13	0.00
Any psychological IPV past 12m	0.03	0.46	0.15	0.00	0.03	0.62	0.02	0.37
Any physical IPV past 12m	-0.01	0.85	0.12	0.00	-0.09	0.11	0.07	0.00
Any sex-related IPV past 12m	0.01	0.80	0.12	0.00	-0.06	0.38	0.02	0.49
Any form of IPV past 12m	0.08	0.09	0.16	0.00	0.05	0.46	0.04	0.19
Frequency of psychological IPV	-1.59	0.28	5.31	0.01	-5.58	0.07	1.39	0.27
Frequency of physical IPV	-0.45	0.28	1.20	0.05	-2.48	0.01	1.24	0.00
Frequency of sex-related IPV	0.27	0.65	1.29	0.07	0.36	0.80	0.47	0.40
Frequency of any IPV	-1.51	0.46	7.86	0.01	-7.83	0.08	3.63	0.07
Marital unrest	-0.04	0.24	0.08	0.02	-0.05	0.36	0.04	0.03
Joint test (*p*-value)		0.42		0.36		0.76		0.03

Note: This table tests for balance in baseline characteristics between arms across the four groups of spousal schooling gap and women's absolute schooling levels.

Appendix Table B11: Violence effects by women's absolute and relative schooling

	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: High Schooling							
Psychological Violence	0.39	0.053	517	0.37	0.096*	444	0.589
		(0.048)			(0.052)		
		[0.297]			[0.075]		
Physical Violence	0.12	0.008	557	0.07	0.121**	481	0.055
		(0.038)			(0.045)		
		[0.763]			[0.020]		
Sexual Violence	0.26	-0.087**	499	0.16	0.111**	432	0.002
		(0.032)			(0.052)		
		[0.020]			[0.051]		
Panel B: Low Schooling							
Psychological Violence	0.45	0.037	238	0.43	0.019	1383	0.901
		(0.142)			(0.035)		
		[0.772]			[0.772]		
Physical Violence	0.27	-0.131	264	0.13	0.053*	1514	0.058
		(0.090)			(0.031)		
		[0.148]			[0.088]		
Sexual Violence	0.36	-0.288***	233	0.26	-0.017	1328	0.000
		(0.069)			(0.035)		
		[0.001]			[0.772]		

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. N reports the effective sample size of each regression. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in column (2) and (5). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf p-values are reported in squared brackets under the SEs.

Appendix Table B12: WHO IPV Items by schooling gap - high schooling women only

	W>H		W≤H		
	C. Mean (1)	Beta (SE) (2)	C. Mean (3)	Beta (SE) (4)	*p*-value (5)
Panel A: Psychological IPV					
Belittled her	0.29	-0.001	0.27	0.134**	0.095
		(0.047)		(0.054)	
Intimidated her	0.16	-0.016	0.16	0.061	0.270
		(0.037)		(0.053)	
Threatened to hurt her	0.04	0.004	0.03	0.001	0.893
		(0.018)		(0.020)	
Called her ugly	0.10	-0.043	0.08	-0.018	0.640
		(0.039)		(0.041)	
Destroyed her belongings	0.03	-0.006	0.03	-0.012	0.825
		(0.017)		(0.017)	
Threatened to take other wife	0.08	-0.070***	0.06	0.013	0.008
		(0.021)		(0.035)	
Threatened to abandon her	0.08	0.015	0.08	0.069	0.258
		(0.025)		(0.048)	
Threatened to divorce her	0.07	-0.029	0.04	0.022	0.186
		(0.028)		(0.029)	
Complains she doesn't please him sexually	0.07	-0.046	0.07	0.006	0.141
		(0.034)		(0.028)	
Panel B: Physical IPV					
Slapped her	0.10	-0.000	0.06	0.168***	0.009
		(0.038)		(0.047)	
Choked her	0.02	0.017	0.02	0.038	0.451
		(0.010)		(0.024)	
Threatened to use a weapon	0.02	0.010	0.02	0.017	0.738
		(0.010)		(0.017)	
Used a weapon	0.02	-0.002	0.02	0.022	0.237
		(0.012)		(0.018)	
Panel C: Sexual IPV					
Forced her to have sex	0.14	-0.020	0.10	0.079	0.128
		(0.034)		(0.052)	
Had sex when did not want	0.20	-0.065*	0.12	0.058	0.026
		(0.037)		(0.045)	
Had degrading sex	0.05	-0.025	0.03	0.020	0.058
		(0.020)		(0.019)	

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. The last column reports the *p*-value of a test of equality between the two estimated treatment effect coefficients reported in columns (2) and (4). Physical IPV includes data from years 1-8. Estimates for psychological and sexual IPV only have data from years 1-7. Robust standard errors clustered at unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table B13: Effects on relationship quality - high schooling women only

	W>H		W≤H		
	Control mean	β (SE)	Control mean	β (SE)	p-value
	(1)	(2)	(3)	(4)	(5)
Relationship z-score	-0.36	0.418	0.02	-0.505**	0.008
		(0.264)		(0.209)	
Any item	0.85	0.120*	0.97	-0.221***	0.003
		(0.067)		(0.074)	
Sum of items	3.92	0.663	4.55	-0.909**	0.029
		(0.545)		(0.375)	
He looks after my needs	0.69	0.104	0.87	-0.201**	0.024
		(0.096)		(0.086)	
He gives me pocket money	0.51	0.048	0.63	-0.052	0.499
		(0.098)		(0.110)	
He understands my feelings	0.67	0.085	0.71	-0.155*	0.095
		(0.110)		(0.082)	
He supports me in difficult sit.	0.72	0.118	0.82	-0.169*	0.039
		(0.085)		(0.084)	
Happy with my husband's behavior	0.69	0.086	0.68	-0.102	0.246
		(0.131)		(0.090)	
Overall good relationship with husband	0.64	0.221*	0.84	-0.231***	0.005
		(0.115)		(0.078)	

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. N reports the effective sample size of each regression. The last column reports the p-value of a test of equality between the two estimated treatment effect coefficients reported in column (2) and (5). Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Romano-Wolf p-values are reported in squared brackets under the SEs.

Appendix Table B14: Characteristics of Peer Volunteers

	Mean (1)	SD (2)	Min (3)	Max (4)	N (5)
Age	29.89	5.71	19	45	45
Un-Married	0.27	0.45	0	1	45
Married	0.69	0.47	0	1	45
Divorced	0.04	0.21	0	1	45
Number of Children	2.07	1.97	0	8	45
Education Years	11.91	2.00	10	16	45
Self Esteem Score	40.65	4.82	30	50	43
Motivation Score	93.58	11.29	49	105	43
Travel time to participants (min)	26.71	20.42	5	90	42
Competency after field training	81.09	7.94	63	95	45

Note: The above table shows summary statistics of the original 45 peer volunteers. Of these, 11 were replaced throughout the intervention.

Appendix Table B15: Do patient characteristics predict PV characteristics?

	Age of IPV (1)	Num. of Children of PV (2)	Schooling Yrs of PV (3)	Married PV (4)	Self-esteem of PV (5)	Motivation of PV (6)	Competency of PV (7)
Mother's age	-0.075	-0.069*	0.022	0.005	-0.041	-0.279	-0.003
	(0.093)	(0.034)	(0.032)	(0.008)	(0.073)	(0.186)	(0.127)
# Children	-0.420	0.178	-0.076	0.018	-0.223	0.084	0.394
	(0.297)	(0.112)	(0.101)	(0.024)	(0.235)	(0.596)	(0.408)
Mother's education	-0.086	-0.024	-0.011	-0.006	-0.166	0.064	0.263
	(0.133)	(0.050)	(0.045)	(0.011)	(0.106)	(0.269)	(0.183)
PHQ-9	-0.231*	-0.057	0.028	0.016	0.057	-0.216	0.100
	(0.114)	(0.040)	(0.039)	(0.009)	(0.090)	(0.229)	(0.156)
WHODAS	0.059	0.016	-0.008	-0.004	-0.034	0.041	-0.014
	(0.049)	(0.017)	(0.017)	(0.004)	(0.040)	(0.101)	(0.068)
PSS Score	-0.028	-0.013	0.016	0.003	-0.048	-0.148	-0.045
	(0.059)	(0.020)	(0.020)	(0.005)	(0.048)	(0.122)	(0.081)
SCID	-0.922	0.412	-0.134	-0.099	-0.776	0.812	1.298
	(0.912)	(0.315)	(0.311)	(0.074)	(0.726)	(1.846)	(1.250)
People per room	0.372	-0.010	-0.056	-0.017	0.081	0.046	-0.079
	(0.193)	(0.064)	(0.066)	(0.016)	(0.152)	(0.386)	(0.265)
SES score	0.138	0.067	0.049	-0.006	0.066	-1.077*	0.214
	(0.266)	(0.095)	(0.091)	(0.022)	(0.211)	(0.537)	(0.365)
Husband's education	0.011	0.074	-0.031	0.002	0.144	0.144	-0.115
	(0.141)	(0.048)	(0.048)	(0.012)	(0.111)	(0.283)	(0.193)
Any IPV in past 12 mnths	1.464	0.577	0.109	-0.093	0.362	-1.414	-0.205
	(0.856)	(0.308)	(0.292)	(0.070)	(0.679)	(1.725)	(1.173)
Frequency of IPV in past 12 mnths	0.239	-0.003	0.046	-0.008	0.258	1.097*	-0.027
	(0.244)	(0.082)	(0.083)	(0.020)	(0.197)	(0.501)	(0.335)
Schooling gap: W > H=2	0.013	-0.122	0.096	-0.069	-1.358	-0.330	-1.470
	(1.154)	(0.408)	(0.393)	(0.094)	(0.918)	(2.334)	(1.582)
Observations	269	198	269	269	262	262	269
p-value (Joint test)	0.027	0.182	0.952	0.468	0.495	0.448	0.837

Note: Regressions include only women (patients) in the treatment arm. Robust standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Appendix Table B16: Treatment effects for W>H by PV characteristics

	Schooling			Competency			Motivation		
	10-12 Yrs (1)	14-16 Yrs (2)	*p*-val (3)	Bottom 50% (4)	Top 50% (5)	*p*-val (6)	Bottom 50% (7)	Top 50% (8)	*p*-val (9)
Depression									
PHQ-9 Score	-2.158***	-2.880***	0.425	-1.501	-2.534***	0.389	-2.158***	-2.818***	0.391
	(0.571)	(0.775)		(1.156)	(0.522)		(0.667)	(0.589)	
PHQ-9 > 10 (0/1)	-0.140***	-0.214***	0.198	-0.167**	-0.164***	0.953	-0.169***	-0.164***	0.910
	(0.032)	(0.053)		(0.063)	(0.033)		(0.040)	(0.041)	
Psychological IPV									
Probability	0.063	-0.014	0.443	0.118	0.024	0.475	0.074	-0.009	0.332
	(0.055)	(0.086)		(0.121)	(0.054)		(0.063)	(0.068)	
Frequency	0.342	-0.520	0.622	3.341	-0.564	0.249	0.779	-0.981	0.407
	(1.656)	(1.741)		(3.269)	(1.570)		(1.707)	(2.007)	
Physical IPV									
Probability	-0.021	-0.046	0.657	-0.026	-0.030	0.938	-0.015	-0.051	0.454
	(0.038)	(0.064)		(0.033)	(0.045)		(0.043)	(0.053)	
Frequency	-0.013	-0.216	0.403	0.105	-0.109	0.409	-0.080	-0.079	0.997
	(0.189)	(0.279)		(0.197)	(0.213)		(0.187)	(0.248)	
Sexual IPV									
Probability	-0.100**	-0.175***	0.132	-0.089	-0.132***	0.532	-0.087**	-0.183***	0.017
	(0.038)	(0.050)		(0.060)	(0.039)		(0.037)	(0.043)	
Frequency	-0.833*	-1.649**	0.221	-0.122	-1.275***	0.112	-1.071*	-1.363**	0.653
	(0.419)	(0.707)		(0.685)	(0.467)		(0.560)	(0.534)	

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. Columns (3) and (6) report the *p*-value of a test of equality between the two estimated treatment effect coefficients reported in columns (1) and (2), as well as (4) and (5), respectively. Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table B17: Treatment effects for W≤H by PV characteristics

	Schooling			Competency			Motivation		
	10-12 Yrs (1)	14-16 Yrs (2)	*p*-val (3)	Bottom 50% (4)	Top 50% (5)	*p*-val (6)	Bottom 50% (7)	Top 50% (8)	*p*-val (9)
Depression									
PHQ-9 Score	-2.202**	-1.179	0.185	-0.938	-2.250**	0.231	-1.276	-2.141***	0.299
	(0.813)	(0.739)		(0.782)	(0.878)		(0.922)	(0.710)	
PHQ-9 > 10 (0/1)	-0.108	-0.053	0.377	-0.047	-0.107	0.444	-0.070	-0.098	0.670
	(0.066)	(0.061)		(0.056)	(0.072)		(0.074)	(0.060)	
Psychological IPV									
Probability	0.096	0.096	1.000	0.028	0.131**	0.311	0.169**	0.040	0.044
	(0.061)	(0.091)		(0.081)	(0.062)		(0.072)	(0.056)	
Frequency	2.814	-0.275	0.253	-1.474	3.381	0.167	-0.011	2.916	0.264
	(2.307)	(2.276)		(2.652)	(2.504)		(2.238)	(2.445)	
Physical IPV									
Probability	0.154**	0.060	0.195	0.109	0.128**	0.873	0.165**	0.086	0.387
	(0.058)	(0.053)		(0.090)	(0.063)		(0.070)	(0.061)	
Frequency	0.424*	-0.085	0.028	0.041	0.326	0.565	0.481	0.060	0.184
	(0.222)	(0.224)		(0.400)	(0.245)		(0.296)	(0.231)	
Sexual IPV									
Probability	0.118*	0.098	0.825	0.182**	0.072	0.293	0.092	0.125*	0.705
	(0.067)	(0.070)		(0.082)	(0.063)		(0.068)	(0.066)	
Frequency	0.601	0.572	0.981	0.769	0.454	0.784	0.610	0.491	0.917
	(0.941)	(0.912)		(0.901)	(0.904)		(1.017)	(0.832)	

Note: $W > H$ indicate women with more schooling than their husbands; $W \leq H$ those with equal or lower schooling. Columns (3) and (6) report the *p*-value of a test of equality between the two estimated treatment effect coefficients reported in columns (1) and (2), as well as (4) and (5), respectively. Robust standard errors clustered at the unit of randomization are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.