Depression, Violence and Bargaining Power

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

September 2, 2026

Abstract

We investigate whether treating maternal depression reduces women's risk of experiencing intimate partner violence (IPV). We leverage variation in depression created by a randomized psychological intervention, following which we tracked both maternal depression and IPV through eight years. We find that treatment reduced maternal depression, but increased women's likelihood of experiencing IPV. These average effects, however, mask divergent trajectories shaped by women's bargaining position in the household, which we proxy with the baseline gap in schooling between spouses. Among women with stronger baseline leverage, psychotherapy reduced sexual violence, increased participation in household decision-making and the labor force, and lowered fertility through greater contraceptive use. Women with lower baseline leverage than their husbands that experienced similar initial reductions in depression, suffered increased psychological, physical, and sexual violence, decreased involvement in household decision-making and lower labor force participation. For this group, the initial mental health gains are reversed. Our results show that psychological interventions can be either undermined or reinforced by the household power dynamics they disrupt.

Keywords: psychological intervention, depression, intimate partner violence, bargaining power, autonomy, fertility, women's labor force participation

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: Business School, Newcastle University (email: michelle.escobar@newcastle.ac.uk); Maselko: Gillings School of Global Public Health, University of North Carolina (email: jmaselko@unc.edu). Sonia Bhalotra acknowledges funding for her time from the European Research Council (ERC) Advanced Grant No. 885698, and from ESRC Centre grant ES/Z504701/1 to CAGE at Warwick.

1 Introduction

Does treating depression reduce women's exposure to intimate partner violence, or could it sometimes increase it? Depression and intimate partner violence (IPV) are among the leading contributors to the global disease burden, frequently co-occurring and affecting hundreds of millions of women worldwide (Vos et al., 2012). Studies in the public health literature have identified depression as a risk factor for IPV victimization (Devries et al., 2013; Oram et al., 2022; Trevillion et al., 2012). It follows that treating a woman's symptoms of depression could reduce her risk of experiencing violence. We provide the first experimental test of this hypothesis using a randomized psychological intervention. We show that while treatment is effective at reducing depression, *on average*, it increases women's likelihood of experiencing intimate partner violence.

To understand the relationship between depression and IPV, consider why husbands use violence in the first place. One view, implicit in the public health literature, is that some husbands, especially those with a taste or preference for violence (Tauchen et al., 1991), exploit the symptoms of depression that impair their partner's capacity for self-advocacy: negative self-beliefs, fatigue, and a diminished sense of self-efficacy (Beck, 1976; Leykin et al., 2011). This view assumes that violence is primarily predatory, and that restoring a woman's psychological functioning is sufficient to reduce her risk. However, economic models of the household recognize that violence can also be instrumental, used to extract resources or compliance (Eswaran & Malhotra, 2011), or triggered by identity threat when the woman's behavior challenges her partner's sense of status or masculinity (Baranov et al., 2021; Cullen et al., 2025).[1] If either of these two motives is operative, treating depression could increase violence if the woman lacks the leverage to back her restored self-advocacy, which determines whether the husband can retaliate against her (Farmer & Tiefenthaler, 1997; Lundberg & Pollak, 1993).[2] If, on the other hand, she can credibly threaten to withdraw cooperation within the marriage — the relevant outside option in a setting where divorce is rare — then violence is costly for the husband and unlikely to escalate.

In general, it is challenging to identify the influence of depression on IPV because common factors cause both, and each causes the other. We address this challenge us-

[1] See Baranov et al. (2021) for a review of how violence has been modeled within the household bargaining framework.

[2] Throughout the paper, we use *leverage* to capture what a woman's bargaining position gives her in practice: the ability to make violence costly for her husband, whether by withdrawing her cooperation in the household or by calling on the support of family and community. In rural Pakistan, Jacoby and Mansuri (2010) find less abuse and better mental health among women whose marriage gives their own family the means to retaliate against the husband. Put simply, a bargaining position is what a woman brings to the marriage; leverage is what she can do with it.

ing a cluster-randomized controlled trial that treated perinatal depression among 570 women in rural Pakistan. The intervention was based on the principles of cognitive behavioral therapy (CBT), the clinical gold standard for mild to moderate depression, delivered by trained peer volunteers from women's own communities (Maselko et al., 2020; Sikander, Ahmad, Atif, et al., 2019). Following randomization, women were surveyed at frequent intervals for eight years, providing rich panel data on the co-evolution of depression, intimate partner violence, and other more cooperative domains of household bargaining. Rates of IPV victimization are 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.

Women's bargaining position at the start of the marriage is proxied using the gap in schooling years between spouses: whether a woman has more education than her husband ($W > H$) or equal or lower education ($W \leq H$). Relative schooling is commonly used to proxy a woman's bargaining position within the household (Doss, 2013; Li, 2026; Rapp et al., 2012). We validate this measure in our data, showing that women with more schooling than their husbands exhibit greater baseline involvement in household decisions, a stronger ability to realize their stated fertility preferences even when spouses disagree, and lower exposure to IPV. Importantly, the spousal education gap is predictive of multiple indicators of women's relative power conditional upon her level of schooling, and the analysis will consistently condition upon the absolute level. Identification of heterogeneity by women who are and are not more educated than their husbands requires that treatment was randomly assigned within each of these groups. We investigate treatment-control balance within group — the joint test across covariates supports this, and we adjust for the imbalance that we observe in a few variables.

Our results can be organized around four sets of findings. The first concerns the intervention's effectiveness at reducing depression: averaging across all eight years of follow-up, treated women score 1.1 points lower on the PHQ-9 depression scale, a 16.5% reduction relative to the control mean, and are 6.7 percentage points less likely to meet clinical thresholds for depression. However, the therapeutic benefits of the intervention are not uniform: treatment effects on depression increase almost monotonically with years of schooling, with no detectable benefit for women with less than four years of education. Among women with secondary schooling or higher, where the therapeutic effects are concentrated, the intervention initially reduces depression in equal measure on both sides of the schooling gap.

Second, turning to violence, we find that assignment to the intervention raises an index of intimate partner violence by 4.2 percentage points on average (8.9% of the

control mean), with physical violence as the primary driver (a 3.5 percentage point, or 25.4%, increase). Results from a list randomization experiment show that these estimates reflect actual increases in the incidence of violence rather than treatment-induced changes in reporting.

Third, this average, masks two sharply divergent patterns, depending on women's initial bargaining position in the household. Among women with more education than their husbands, treatment reduces sexual violence by 8.7 percentage points, a 33% decline relative to the control mean, with no significant change in other forms of violence. Among women with equal or lower schooling than their husbands, treatment increases physical violence by 12.1 percentage points, nearly doubling the control mean, and sexual violence by 11.1 percentage points, a 69% increase.

This asymmetry may seem surprising. If treatment restores a woman's agency, one might expect any backlash to be most severe against women with leverage, whose gains pose the greatest challenge to their husbands. However, bargaining theory predicts the opposite: violence is used where it is less costly to the perpetrator. Where a woman holds initial leverage, her restored self-advocacy is backed by a credible threat to withdraw the surplus she contributes, making escalation costly for her husband. Where she does not have that initial leverage, she recovers her voice with nothing to back it and, in this case, escalation costs her husband little. Control group trajectories support this reading: absent treatment, the baseline advantage of women with leverage inverts, and they come to experience *more* violence than their counterparts.[3] Treatment prevents this decline: rather than provoking new conflict, it restores their bargaining power.

Fourth, these divergent patterns extend beyond violence to other, more cooperative domains of household bargaining. Among women with more education than their husbands, treatment leads to greater involvement in household decisions (0.28 SD), higher labor force participation (6.2 percentage points), and reductions in cumulative fertility achieved through increased use of modern contraceptives (14.4 percentage points).[4] Among women with equal or lower schooling, treatment reduces involvement in household decision-making (0.21 SD) and lowers labor force participation (6.9 percentage points), with no significant effect on fertility.

These broader patterns are informative in two respects. First, the similarity of the

[3]Untreated women with more education than their husbands also come to experience more violence than comparable non-depressed women in our reference cohort, suggesting that the escalation reflects untreated depression rather than the bargaining position itself.

[4]The fertility result is informative about women's bargaining power because spouses disagree over fertility. Husbands desire more children than wives in this setting, consistent with the gap documented across developing countries (Doepke & Tertilt, 2018). When preferences diverge in this way, contraceptive use responds to women's bargaining power (Ashraf et al., 2014), so the decline in fertility suggests that women's preferences carry greater weight in household decisions.

results across domains provides independent validation of the violence findings and supports our suggestion that the underlying mechanism is that CBT triggers women's self-advocacy: improvements across the board for women with baseline leverage, deterioration across the board for women without it. This is not a pattern that can be explained by differential reporting of sensitive questions alone. Second, the reductions in decision-making and labor force participation among the women with lower baseline leverage shed some light on the motives behind the violence. If violence were purely instrumental, we would expect women to return to control levels of decision-making and economic participation once compliance is achieved. The fact that treated women fall *below* the control group is harder to reconcile with a purely instrumental motive. While we cannot separate instrumental from identity-based motives, we interpret this over-correction as suggestive evidence that identity threat operates alongside instrumental motives: the husband may be punishing the change itself, not merely reasserting control.

These results demonstrate how intrahousehold responses can undermine interventions designed to improve women's wellbeing. Recent experimental evidence shows that legal reforms granting women access to driving and employment produce similarly divergent responses (Daher et al., 2025). Whether the intervention operates through psychological empowerment (using psychological intervention) or through economic empowerment (such as improving women's legal rights or using economic transfers), the mechanism is similar: gains in women's agency are mediated by the household. This does not mean that treating depression necessarily harms women: potential harm arises where violence carries little cost for the husband (weakly sanctioned, with exit from marriage rare) and where treatment gives women the motivation to renegotiate the terms of the marriage without the leverage to back this. The risk lies in scaling psychosocial programs blind to this distinction.

This paper makes three contributions. First, we contribute to the literature on the economics of intimate partner violence by producing the first identified estimates of how mental health can potentially shape self-advocacy and preference articulation. We highlight the complementarity between mental health and baseline leverage. If depression numbs the capacity or motivation for self-advocacy, leverage remains latent. However, if depression treatment triggers self-advocacy the extent to which this is effective depends on baseline leverage. This insight helps explain heterogeneous responses to empowerment-oriented interventions and extends beyond violence and households. Related work documents that interventions aimed at improving women's economic position can have heterogeneous effects on IPV.[5]

[5]Some studies find that improvements in women's outside options reduce violence, including divorce law reforms (Stevenson & Wolfers, 2006), income gains and transfers (Angelucci, 2008; Bobonis

Second, we contribute to a growing public health literature on the bi-directional relationship between intimate partner violence and mental health disorders (Oram et al., 2022; Trevillion et al., 2012). To the best of our knowledge, this is the first study to document adverse effects of depression treatment on future IPV victimization for a general population, albeit our focus is on maternal depression. Our intervention does not include IPV-specific content, allowing us to isolate the effects of treating mental health symptoms.[6] Our findings for women with more education than their husbands demonstrate that, where women are able to assert the agency gained from psychological treatment, such treatment can result in a decline in IPV in a more general setting.

Lastly, we contribute to a growing literature in economics documenting the long-run effects of cognitive behavioral therapy-based interventions (Baranov et al., 2020; Blattman et al., 2023) by providing continuous evidence on the trajectory of treatment effects over eight years. Among women whose overall wellbeing and empowerment improve as a result of the intervention, the initial mental health gains are fully sustained. Among women who experience escalating violence, the initial gains in mental health erode progressively, and in periods where violence is most acute, treated women report higher depression scores than the control group. This pattern is consistent with a feedback loop in which violence becomes a source of renewed depressive symptoms, unwinding the benefits of treatment. Thus, whether mental health gains persist depends on how the household responds to the woman's improved wellbeing.

The rest of the paper proceeds as follows. Section 2 outlines our conceptual framework. Section 3 describes the experimental design. Section 4 describes the data and Section 5 outlines our empirical strategy. Section 6 presents average treatment effects on depression and IPV. Section 7 focuses on the role of women's leverage. Section 8 explores alternative mechanisms, and Section 9 concludes.

2 Conceptual Framework

In the bargaining framework often used in the economics literature, whether a husband engages in intimate partner violence is jointly determined by his motives for using violence and by the constraints his wife can impose on that choice (Anderberg

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 (Cullen et al., 2025; Daher et al., 2025; Erten & Keskin, 2018, 2024; Guarnieri & Rainer, 2021; Roy et al., 2019). Standard economic models rationalize these mixed findings by allowing violence to serve different functions (Eswaran & Malhotra, 2011; Farmer & Tiefenthaler, 1997; Tauchen et al., 1991).

[6]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 (Spencer et al., 2022; Tol et al., 2019).

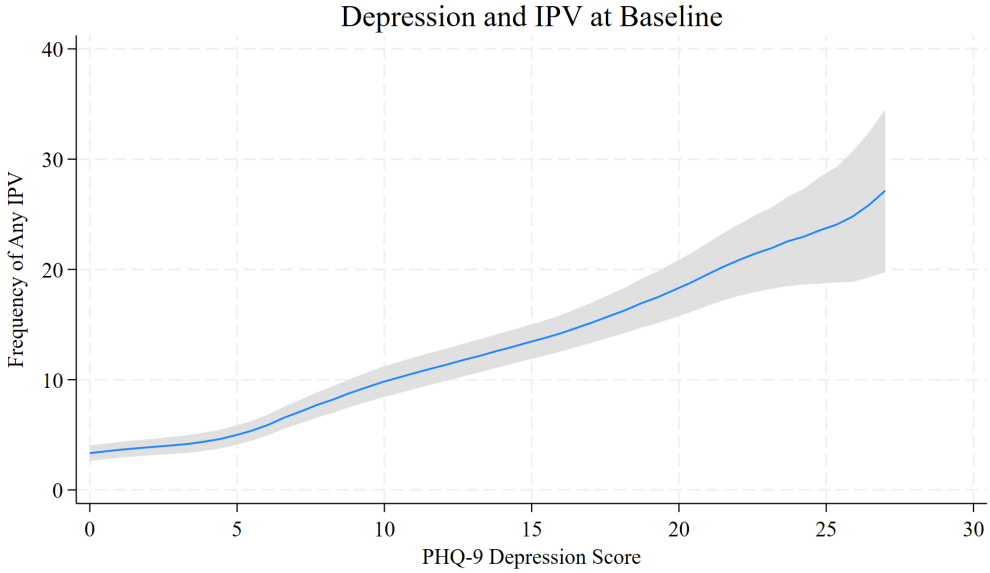
et al., 2016). The literature distinguishes three broad motives. First, a man may derive direct utility from violence – what the literature terms expressive or taste-based violence (Tauchen et al., 1991). Second, violence may be instrumental: a strategic tool deployed to extract compliance, control household resources, or maintain authority within the relationship (Anderberg & Rainer, 2013; Bloch & Rao, 2002; Eswaran & Malhotra, 2011). Third, violence may be triggered by identity threat: changes in a woman's behavior that challenge her partner's sense of status or masculinity can provoke a violent response aimed not at restoring a specific allocation but at punishing the deviation itself (Cullen et al., 2025). Each of these motives generates a distinct prediction about what happens when a woman's depression is treated.

Consider first the taste-based channel. When violence reflects a man's preferences or predispositions, its incidence depends on his partner's ability to resist, seek help, or mobilize social and familial support (Farmer & Tiefenthaler, 1997; Tauchen et al., 1991). Depression erodes precisely these capacities. Clinical evidence documents that depression impairs problem-solving ability, reduces initiative, and narrows the range of coping strategies available to affected individuals (Joormann & Gotlib, 2010; Nezu, 1986). In the context of intimate partner violence, these impairments translate into diminished capacity to resist abuse, recognize escalating danger, or act on available exit options. From this perspective, the robust positive cross-sectional association between depression and IPV victimization, visible in our own data (Figure 1), reflects the exploitation of psychological vulnerability. The policy implication is clear: treating depression could restore women's capacity for self-protection, thereby reducing violence.

When violence is instrumental, it is employed not because it is valued per se but because it alters the bargaining outcome – extracting a larger share of household resources, enforcing compliance with the husband's preferred allocation, or deterring the wife from pursuing activities he opposes (Bloch & Rao, 2002; Eswaran & Malhotra, 2011). In this framework, a depressed woman who is passive, withdrawn, and unable to articulate her preferences poses little threat to her husband's preferred allocation. Treating her depression restores her capacity for self-advocacy and preference articulation – she may begin to voice disagreement over household spending, resist demands she previously accepted, or seek greater involvement in decisions from which she was excluded. These are precisely the behaviors that instrumental violence is designed to suppress. Treatment may therefore trigger the very violence it was expected to prevent, as the husband deploys violence to curtail the newly expressed advocacy.

The identity-threat channel yields a similar prediction through a different mechanism. Here, violence is triggered not by a strategic calculation about household allocations but by a perceived challenge to the husband's sense of status, authority, or

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.

masculinity (Cullen et al., 2025). When a woman who was previously compliant and withdrawn becomes more assertive, more vocal, or less deferential as her depression lifts, her changed behavior may be experienced by her partner as a threat to his identity within the household and the broader community. The resulting violence is reactive rather than strategic: its purpose is to punish the deviation from expected behavior, not to extract a specific concession. This distinction matters empirically. If violence were purely instrumental, we would expect it to cease once the woman returns to the husband's preferred level of compliance. If identity threat is also at play, violence may go beyond what is needed to restore the status quo.

Which of these forces dominates in practice depends on the constraints the woman can impose on her partner's use of violence, what we refer to as her relative standing, or potential bargaining power, within the household. Classical bargaining models emphasize the role of outside options: the credibility of a woman's threat to withdraw cooperation or exit the relationship (Bhalotra et al., 2021; Farmer & Tiefenthaler, 1997; Lundberg & Pollak, 1993). Where a woman holds sufficient standing, the expected costs of violence to the perpetrator are high and the protective effects of reduced vulnerability are likely to dominate. Where she does not, the increases in self-advocacy that accompany treatment may trigger escalation rather than protection. In our setting, as in much of rural South Asia, divorce is permitted under Islamic law but carries deep social stigma and is rarely pursued (Bloch & Rao, 2002). Exit from marriage is therefore not a credible threat for most women, and standing instead operates through

other channels: the ability to withdraw cooperation within the household, to mobilize support from natal kin, or to draw on social and economic resources that raise the cost of violence to the husband. Anderson and Eswaran (2009) reach a similar conclusion in Bangladesh, where women's autonomy is better explained by what they can do within the marriage than by the option of leaving it.

This framework yields two testable predictions. First, the net effect of treating depression on violence should vary with women's relative standing in the household. Among women with sufficient standing to impose costs on a violent partner, reductions in depression should decrease violence as the protective effects of reduced vulnerability dominate. Among women who lack such standing, treatment may increase violence as the instrumental and identity-threat channels operate with few constraints. Second, the framework generates predictions beyond violence. Among women with higher standing, if reductions in violence reflect a broader shift toward more cooperative relationships, we would also expect improvements in other domains: greater involvement in household decisions, higher labor force participation, and fertility outcomes that better reflect women's stated preferences. Among women with lower standing, what happens to these cooperative domains depends on the motive behind the violence. If violence is purely instrumental, in this case aimed at restoring compliance, we would expect treated women to return to control levels of decision-making and economic participation once the husband's objective is achieved. That is, there should be no effect on these outcomes. If identity threat is also at play, treated women may fall below control levels in these domains, as the husband punishes the change rather than simply restoring the status quo.

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 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). 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).[7]

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 25.3% and by year 8, 73% 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 B5, 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

[7] 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 B3 - B7 show balance tests across all follow-up waves by schooling gap group.

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

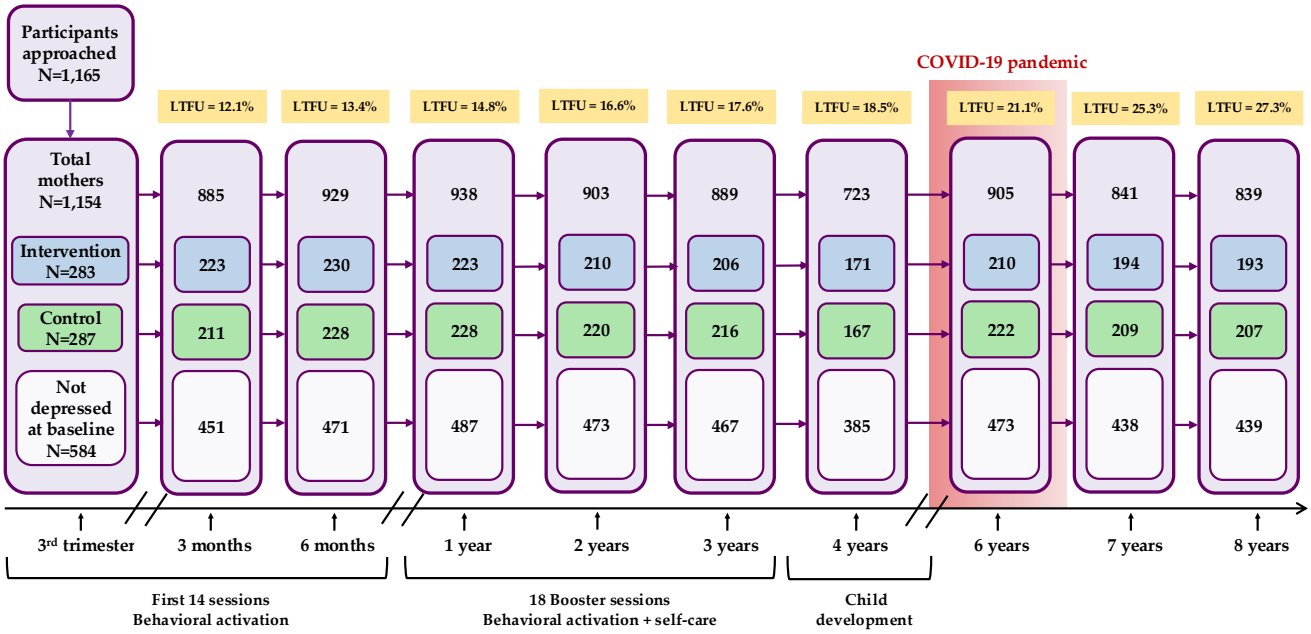
Note: This table tests for balance in baseline characteristics. Columns 1 and 3 show the mean in the control, and treatment group at baseline.

(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).[8] Additionally, in year 8 of follow-up, we measured the experience of

[8]See Appendix Table B8 for a complete list of all the items.

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. A respondent is coded as lost to follow-up (LTFU) after the last one at which she was observed. The indicator is monotonic within woman, so intermittent missingness between two observed waves is treated as in-sample rather than as attrition.

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", "11-20 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

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 de-

cisions 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.

Fertility and Contraceptive Use.— 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). Lastly, we ask women about use of various forms of contraception. Here, we focus on modern contraceptives as the outcome of interest: the pill, IUDs, etc.

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 specification.[9] First we report the impact of treatment at each follow-up wave by interacting the treatment indicator with an indicator for the wave:

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

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.

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

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.[10] 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.

6 Treatment Effects on Depression and IPV

We first show average treatment effects on maternal depression and intimate partner violence, pooling across all post-treatment waves. We then examine the trajectory of these effects over the eight years of follow-up.

A. Effects on Depression

As shown in Table 2, panel A, the intervention produced large and statistically significant reductions in depression across all three measures. Treated women score 1.12 points lower on the PHQ-9 depression scale relative to a control mean of 7.0 (column 1, $p < 0.01$), a 16 percent reduction. The probability of meeting the clinical threshold for moderate-to-severe depression (PHQ-9 > 10) falls by 6.7 percentage points, a 22 percent decline relative to the control mean of 30 percent (column 2, $p < 0.01$). The SCID clinical diagnosis shows a similar pattern: treatment reduces the probability of a major depressive episode by 5.5 percentage points from a base of 20 percent, a 28 percent reduction (column 3, $p < 0.01$).

Figure A1 (top row) and Appendix Table B9 present the wave-by-wave treatment effects underlying these pooled estimates. The trajectory reveals two notable features. First, the largest effects appear immediately: at three months post-partum, treated

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

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 $\geq$ 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.

women score 2.3 points lower on the PHQ-9 ($p < 0.01$) and are 13 percentage points less likely to screen positive ($p < 0.01$). Second, while the point estimates attenuate somewhat in the medium run, the effects re-emerge at the seven- and eight-year follow-ups, where treated women again score significantly lower on the PHQ-9 (-1.23, $p < 0.05$ at year 7; -1.19, $p < 0.05$ at year 8). This pattern of persistent effects is consistent with the broader evidence on CBT documenting lasting benefits well beyond the treatment period (Baranov et al., 2020).

We also find that the effects of treatment assignment on depression are monotonically increasing with respondents' schooling, as shown in Figure A2. We are unaware of earlier evidence that depression treatment is more effective at higher levels of schooling. We find no impact of treatment for women with less than four years of schooling.

B. Effects on Intimate Partner Violence

Panel B of Table 2 reports average treatment effects on violence. Despite causing average improvements in mental health, we see that treatment increases women's experience of intimate partner violence. An index of any form of IPV rises by 4.2 percentage

points, an 8.9 percent increase relative to the control mean of 47 percent (column 1, $p < 0.10$). This increase is driven by psychological and physical violence: the probability of experiencing psychological violence increases by 4.1 percentage points, a 10 percent increase relative to the control mean of 41 percent (column 2, $p < 0.10$), and physical violence by 3.5 percentage points, a 27 percent increase relative to the control mean of 13 percent (column 3, $p < 0.10$). We find no significant effect on sexual violence (column 4, $\beta = -0.014$, $p = 0.44$).

The trajectory of violence effects, shown in Figure A1 (bottom row) and Appendix Table B10, differs markedly from that for depression. Treatment effects on violence emerge gradually. At the one-year follow-up, the point estimates on all forms of IPV are small and statistically insignificant. By year two, the coefficient on physical violence reaches 7.1 percentage points ($p < 0.10$); by years seven and eight, the effects on psychological IPV are 8.8 percentage points ($p < 0.10$) and physical IPV is 9.5 percentage points ($p < 0.10$), respectively. This delayed onset is consistent with violence responding to behavioral changes induced by treatment (changes in self-advocacy and household participation that take time to manifest and to provoke a response), rather than to the reduction in depression.

Because our measures of violence are self-reported, a natural concern is that treatment may have increased women's willingness to disclose violence rather than its actual incidence. We address this concern using a list randomization experiment conducted at the six-year follow-up, in which respondents were randomly assigned to receive a list of either four or five statements, with the additional statement concerning whether the respondent had been slapped by her husband. Appendix Table B11 reports the results. The list experiment yields a prevalence estimate of 28 percent among treated women, compared to 15.5 percent under direct questioning (column 2), indicating that if anything, treated women *under report* violence relative to the list experiment benchmark. Among control women, the list and direct estimates are similar (7 percent and 8.9 percent, respectively; column 3). These results indicate that the observed treatment effects on violence reflect genuine increases in incidence rather than changes in reporting.[11]

7 The Role of Women's Leverage

In this section we first validate our proxy for women's relative status at baseline, and then show how treatment effects on depression, violence, and the cooperative domains of household bargaining diverge by it.

[11] The full analysis of the list experiment is reported in Bates et al, mimeo.

A. Measuring Baseline Leverage

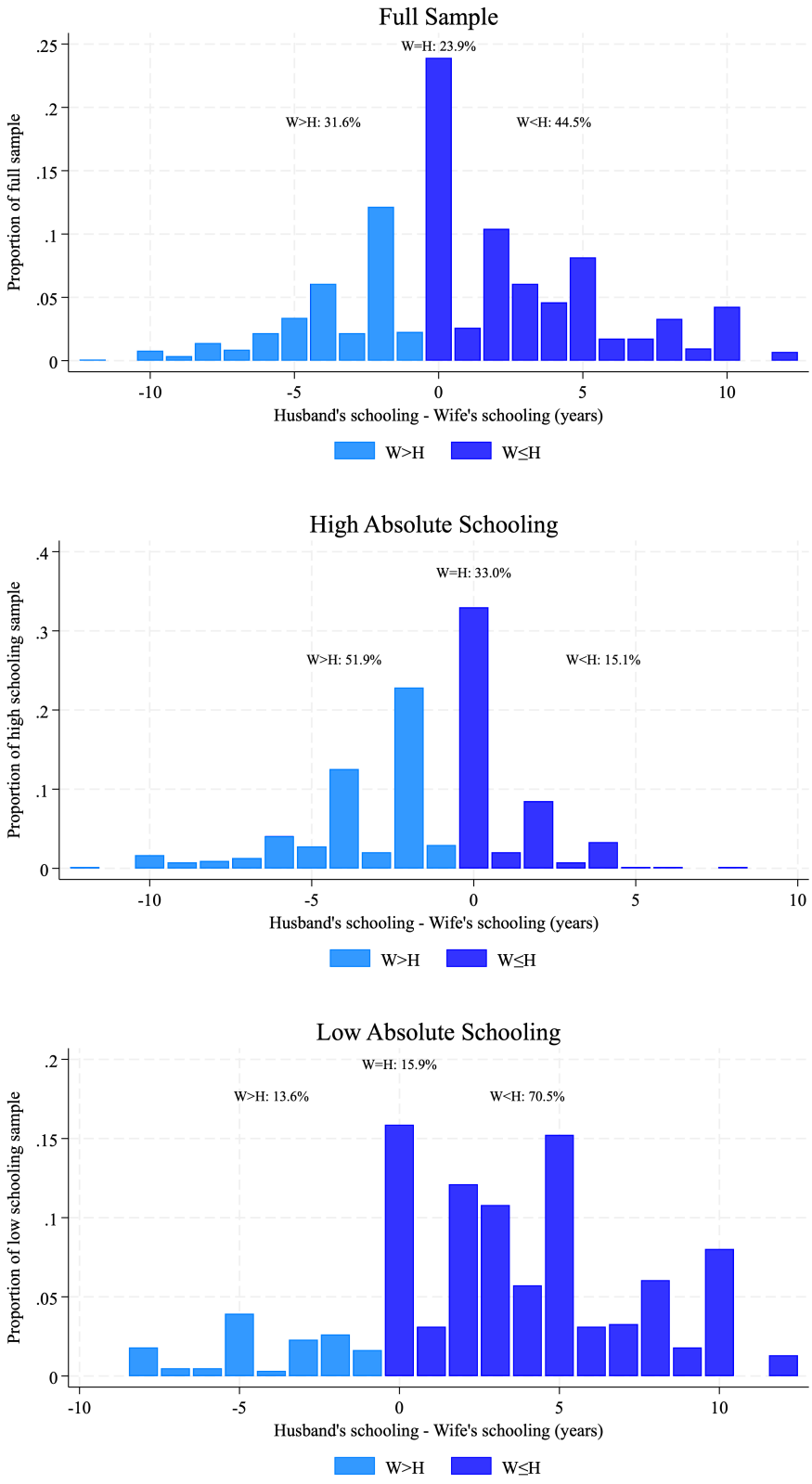
We proxy for a woman's relative status using the gap in schooling years between spouses, distinguishing women with more education than their husbands ($W > H$) from women with equal or lower education ($W \leq H$). Figure 3 plots the distribution of this variable. In the full sample, as expected in a setting where hypergamy is the norm, most women married up or matched their husbands, but a substantial minority of 31.6 percent married down. The lower two panels split the distribution by women's absolute schooling and preview a feature of the data that shapes our empirical approach: among women with secondary schooling or more, marrying down is the majority outcome at 51.9 percent, while among women with low schooling it is rare at 13.6 percent, with 70.5 percent of these women married to more educated men. Relative status and absolute human capital are therefore entangled, a point we return to below.

Three features of the data support the interpretation of this gap as a measure of latent bargaining power. First, the gap predicts women's involvement in household decisions: women with more schooling than their husbands score 0.19 standard deviations higher on the decision-making index and are significantly more likely to have a say in their own healthcare and in major purchases, a pattern we replicate in the rural Pakistan sample of the Demographic and Health Survey (Appendix Table B14).[12] These correlations are robust to controlling for whether each spouse has secondary schooling, so the gap is not simply repackaging absolute education, and they are not driven by depression status. Second, the gap predicts whose fertility preferences bind.[13] Appendix Table B15 shows that a husband's desire for more children predicts realized fertility only among women with equal or lower schooling. Among women who married down, the husband's excess demand for children has no purchase on outcomes, while the woman's own desired fertility predicts realized fertility in both groups. High-status women, in other words, can realize their preferences even when spouses disagree. Third, these women are better off at baseline across virtually every domain we measure: they have greater mobility, stronger support from friends, and experience psychological violence markedly less frequently (Appendix Table B12). Among the reference cohort of women who were not prenatally depressed, violence is indistinguishable across the schooling gap (Figure A5). Neither comparison suggests that marrying down by itself invites retaliation. The pattern is instead consistent with the economic logic laid out in the introduction: a wife who contributes more to household

[12] Beegle et al. (2001) similarly find that Indonesian women who are better educated than their husbands have more say in decisions about their own health care.

[13] Rasul (2008) shows that which spouse's fertility preferences prevail reflects the balance of bargaining power between them.

Figure 3: Distribution of baseline schooling gap between spouses



Notes: Figures plot 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.

surplus than her husband can credibly threaten to withdraw more — and this raises the cost of mistreating her.

Because schooling is also a productive attribute in its own right, the entanglement visible in Figure 3 means the two margins must be separated: fifty-three percent of women with $W \leq H$ have low absolute schooling, compared to nine percent of women with $W > H$. We address this in two ways. First, we condition on women's absolute level of education throughout. Second, and more importantly, the analysis in this section is restricted to women with secondary schooling or more, the group in which the therapeutic effects of the intervention are concentrated. Within this group, women on both sides of the schooling gap benefited equally from treatment, which allows us to hold the first stage fixed and isolate how the household responds to a comparable improvement in the woman's mental health.

A related concern is selection into marriage. Women who married down may differ from women who married up in ways that shape their response to treatment. Two points bear on this. First, marriages in this setting are arranged and contracted early, often before a woman's adult traits are fully revealed, which limits the scope for sorting on unobserved assertiveness or temperament. Second, our identification does not require that the schooling gap be randomly assigned. It requires that treatment was randomly assigned within each gap group. Appendix Table B1 reports baseline balance across arms within each of the four schooling groups, and Appendix Tables B3 and B4 report balance within the two high schooling groups at every follow-up wave, with Appendix Tables B6 and B7 showing that attritor characteristics are likewise balanced across arms within these groups in all waves except year eight. The imbalances that appear in our analysis sample of women with high schooling are modest, and we control for imbalanced baseline characteristics interacted with wave indicators throughout. The threat is therefore not confounding of the treatment effect but mislabeling of the moderator, and the validation exercises above support the bargaining interpretation.

B. Treatment Effects on Depression by Leverage

Table 3 reports treatment effects on depression by absolute schooling (panels) and relative status (columns). Two results stand out. First, the therapeutic effects are concentrated among women with high absolute schooling. In Panel A, treatment reduces PHQ-9 scores by 2.1 points for women with more schooling than their husbands and by 1.8 points for women with equal or less, reductions of roughly 30 percent relative to control means, with corresponding declines in clinical depression on both the PHQ-9 threshold and the SCID diagnosis. All six estimates survive Romano-Wolf adjustment. Among the majority of low schooling women in Panel B, those with equal

or lower schooling than their husbands, the point estimate on the PHQ-9 is precisely zero.[14] Overall, among women who have low absolute education, most marry up and, in this group, there is no depression reduction. The gradient in education is consistent with the cognitive demands of CBT, which asks participants to identify, articulate, and restructure unhealthy thoughts, and with the higher session attendance of more educated women (Figure A2).

Second, and central to our design, the mental health gains among high schooling women are statistically indistinguishable across the spousal gap. The test of equality yields $p = 0.75$ for the PHQ-9 score, and the raw trajectories in Figure 4 show that both groups experience the same steep drop in symptoms in the first six months after treatment, with the initial declines, if anything, somewhat larger among lower-status women. The corresponding treatment effects with confidence intervals are shown in Appendix Figure A3. Where the two groups part ways is not in the mental health response to treatment but in what follows from it. We return to the long-run divergence in these trajectories in subsection E, once the violence results make its source clear.

C. Treatment Effects on Violence by Leverage

Table 4 presents the central result of the paper. Among high schooling women with more education than their husbands, treatment reduces sexual violence by 8.7 percentage points, a 33 percent decline relative to the control mean, with small and insignificant effects on psychological and physical violence. Among high schooling women with equal or lower education than their husbands, treatment increases psychological violence by 9.6 percentage points, physical violence by 12.1 percentage points from a control mean of just 7 percent, and sexual violence by 11.1 percentage points, a 69 percent increase. The tests of equality reject a common treatment effect for physical violence ($p = 0.055$) and decisively so for sexual violence ($p = 0.002$).

The intensive margin points in the same direction. Appendix Figure A6 plots the distribution of the cumulative number of incidents of each form of violence over the follow-up period. Among low-status women, the distribution for treated women lies to the right of the control distribution for all three forms: treatment raises not only whether women experience violence but how often. Among high-status women, the ordering reverses for physical and sexual violence, with treated women concentrated at fewer incidents. The item-level results in Appendix Table B17 add texture: the in-

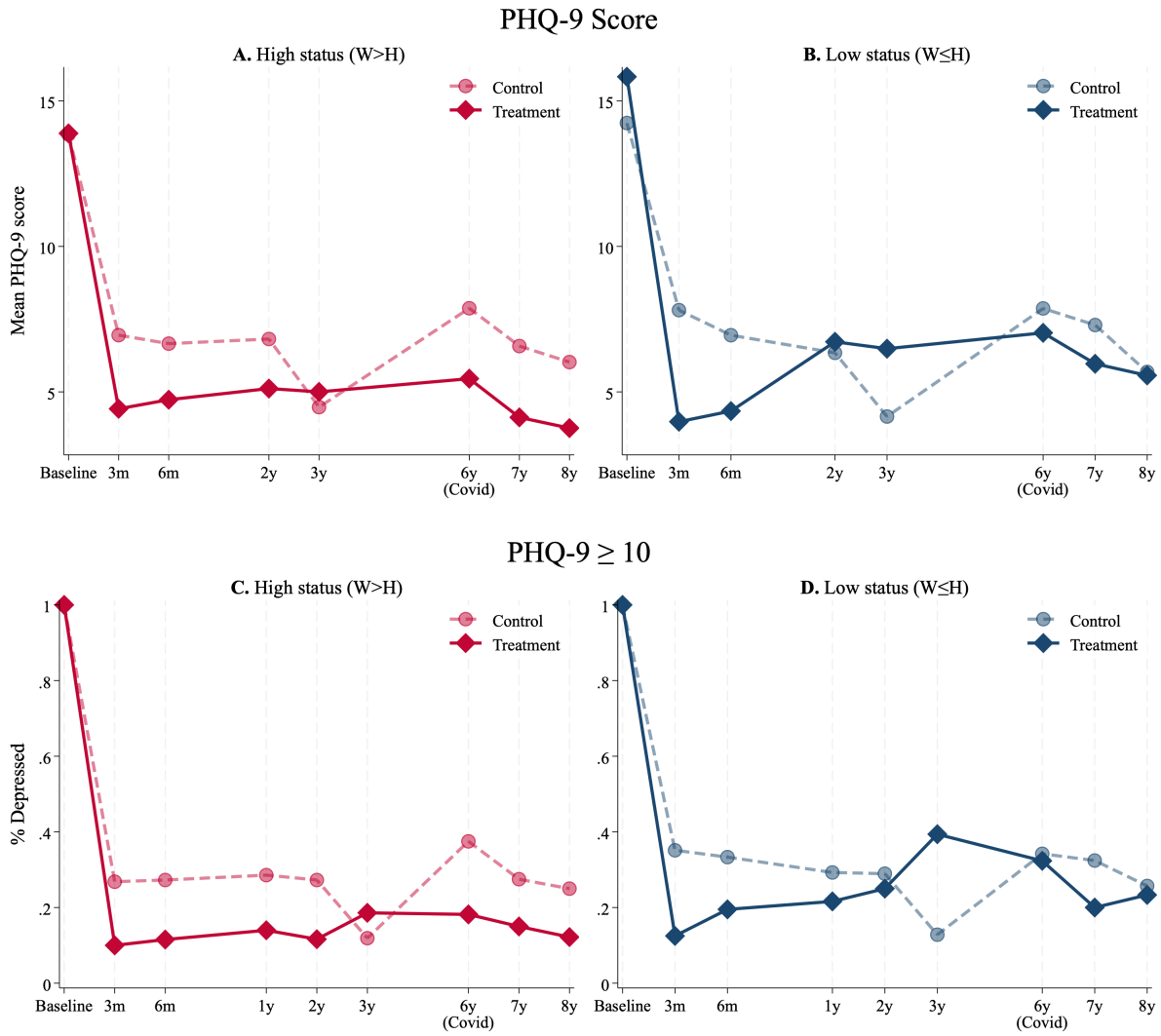
[14] The small group of low schooling women who married up shows mental health benefits and, in Appendix Table B16, a large reduction in sexual violence, a pattern consistent with our framework. We do not emphasize these estimates given the group's size (nine percent of low schooling women). However, this result confirms that the impact of treatment on depression varies with the spousal gap more clearly than it does with the absolute level of education.

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 $\geq$ 10 (0/1)	0.26	-0.154***	816	0.29	-0.105**	682	0.425
		(0.032)			(0.051)		
		[0.001]			[0.002]		
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)		
		[0.120]			[0.919]		
PHQ-9 $\geq$ 10 (0/1)	0.49	-0.145*	361	0.29	0.011	2141	0.053
		(0.080)			(0.025)		
		[0.079]			[0.860]		
SCID (0/1)	0.38	-0.143**	360	0.18	0.006	2140	0.029
		(0.066)			(0.020)		
		[0.049]			[0.919]		

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.

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.

creases among low-status women are concentrated in belittling (13.4 percentage points) and slapping (16.8 percentage points), the classic acts of intrahousehold discipline, while the reductions among high-status women are concentrated in threats to take another wife and coerced sex, acts that exploit a woman's inability to push back.

The trajectories in Figure 5 reveal how these effects unfold, with the corresponding treatment effects and confidence intervals in Appendix Figure A4. Consistent with the delayed onset documented in Section 6, the divergence emerges gradually: physical violence among treated low-status women is already elevated by year two and remains above control levels through year seven, with the gap closing only at year eight when violence in the control group rises, while the protection enjoyed by high-status women

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

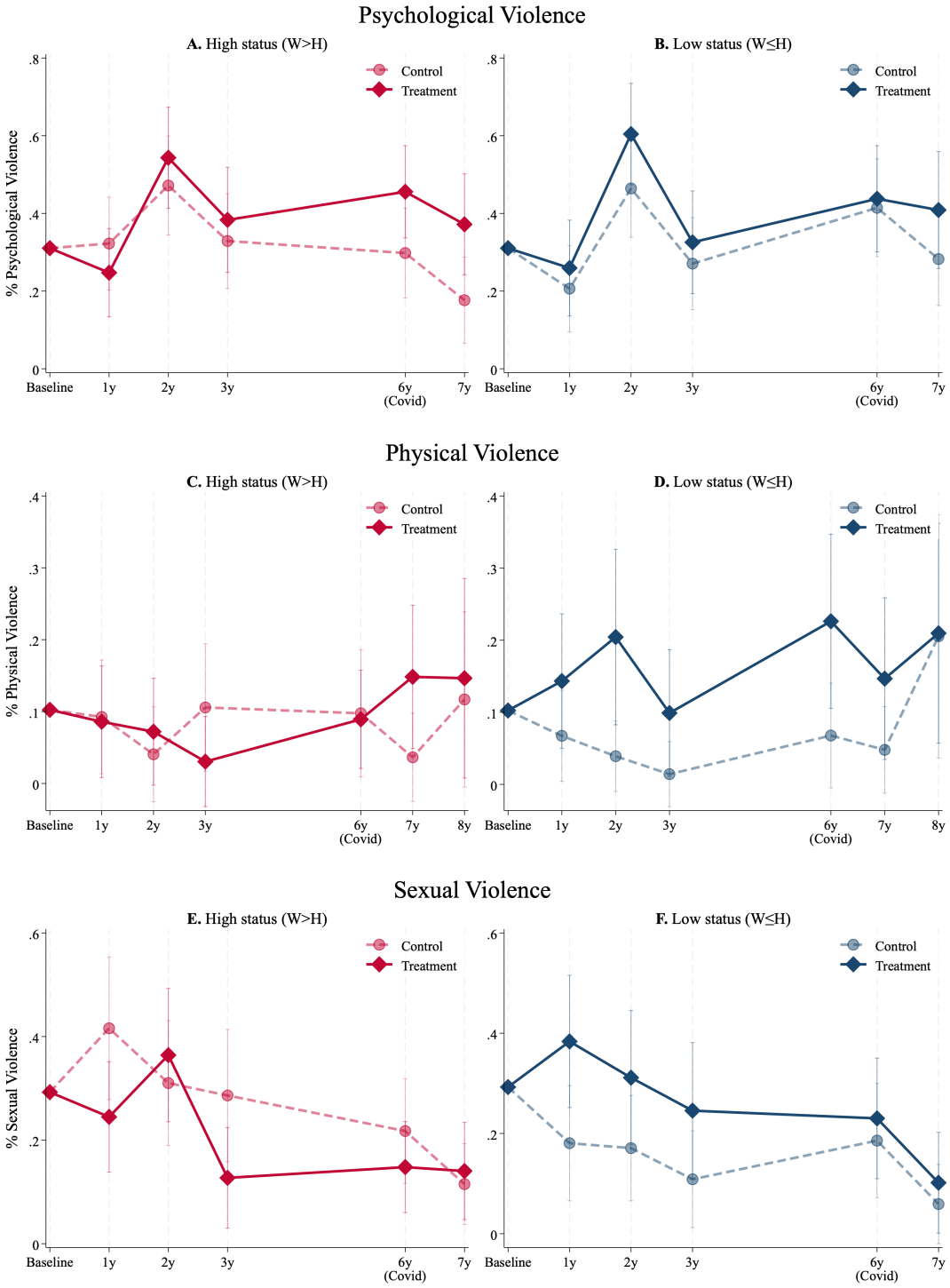
	W>H			W≤H			
	Control mean	β (SE)	N	Control mean	β (SE)	N	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Psychological Violence	0.39	0.053	517	0.37	0.096*	444	0.589
		(0.048)			(0.052)		
		[0.297]			[0.075]		
Physical Violence	0.12	0.008	557	0.07	0.121**	481	0.055
		(0.038)			(0.045)		
		[0.763]			[0.020]		
Sexual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

builds from year two onward. This timing is what we would expect if violence responds to treatment-induced changes in behavior rather than directly to the lifting of depression.

The control group provides an instructive counterfactual. Among untreated depressed women, the baseline position of high-status women erodes and then inverts: post-baseline, they experience more violence than their low-status counterparts on every form we measure, including sexual violence at 26 percent against 16 percent (Appendix Table B13). Figure A5 shows that they also come to experience more violence than non-depressed women of the same status, while for low-status women the depressed control and non-depressed trajectories track each other closely until late in the follow-up. Untreated depression, in other words, is particularly costly for women who have status to lose. A plausible reading is that prolonged depression steadily degrades a woman's capacity to exercise the leverage her status affords. Treatment, on this reading, does not hand high-status women new power. It restores their access to power they already held. For low-status women the same restoration of voice arrives without leverage to back it, and the household response is escalation of violence against her.

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.

D. Beyond Violence: Decision-Making, Work, and Fertility

If the violence results reflect a genuine renegotiation of household arrangements, the divergence should not be confined to violence. Table 5 examines fertility, decision-

making, and labor force participation among women with secondary schooling.

The pattern is strikingly similar. Among women with higher status, treatment increases the use of modern contraceptives by 14.4 percentage points on a control mean of 27 percent, and by year eight treated women have had 0.37 fewer children since the index child, a 30 percent reduction. They score 0.29 standard deviations higher on the decision-making index, are 12 percentage points more likely to be involved in all decision domains, and are 6.2 percentage points more likely to have worked in the last year, a 62 percent increase on a control mean of 10 percent. Among women with lower status, treatment reduces the decision-making index by 0.23 standard deviations and lowers labor force participation by 6.9 percentage points, wiping out most of the control mean, with no detectable effect on contraceptive use or fertility. The tests of equality reject a common effect for decision-making ($p = 0.001$) and labor force participation ($p = 0.002$).

The same divergence appears in women's assessments of the marriage itself. Treated high-status women are 22 percentage points more likely to describe their relationship with their husband as good overall, while treated low-status women report a 0.44 standard deviation decline in the relationship quality index, with significant deterioration in whether the husband looks after their needs, understands their feelings, and supports them in difficult situations (Appendix Table B18).

This symmetry is informative in two respects. First, it provides independent validation of the violence findings. Improvements across the board for one group and deterioration across the board for the other is not a pattern that differential reporting of sensitive questions can easily produce. A disclosure account would require treatment to make low-status women more forthcoming and high-status women less so, and to do so in domains like labor force participation and contraceptive use where reporting incentives are entirely different. Second, the direction of the effects for low-status women speaks to the motives behind the violence. If violence were purely instrumental, aimed at restoring the husband's preferred allocation, treated women should return to control levels of decision-making and economic participation once compliance is achieved. Instead they fall below the control group. This over-correction is difficult to reconcile with a purely instrumental account and is consistent with identity threat operating alongside it. The husband appears to be punishing the change itself, not merely reasserting a preferred allocation.

E. The Erosion of Mental Health Gains

We close this section by returning to the long-run depression trajectories in Figure 4, which now admit a clear interpretation. Among high-status women, the mental health

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							
Uses modern contraceptives	0.27	0.144**	147	0.46	-0.049	127	0.220
		(0.070)			(0.121)		
		[0.243]			[0.954]		
Pregnant	0.10	-0.033	413	0.09	0.029	349	0.107
		(0.026)			(0.023)		
		[0.243]			[0.954]		
# Children alive since index (Yr 8)	1.25	-0.374*	80	1.24	-0.175	67	0.541
		(0.209)			(0.234)		
		[0.709]			[0.709]		
Panel B: Decision-Making							
Decision-making index	-0.03	0.293***	348	0.06	-0.227*	296	0.001
		(0.092)			(0.127)		
		[0.016]			[0.110]		
Woman is involved in all	0.24	0.120***	348	0.26	-0.036	296	0.004
		(0.038)			(0.032)		
		[0.016]			[0.267]		
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.

gains are sustained throughout: treated women report lower depression than controls in every wave, with the largest gap at the six-year follow-up, when depression in the control group spikes during the COVID-19 pandemic and treated women are largely protected. The re-emergence of significant pooled treatment effects at years seven and eight documented in Section 6 is driven by this group.

Among low-status women, the initial gains erode as violence escalates. Treated women converge to control levels of depression by year two, and at year three, the wave following the sharp rise in physical violence, they report depression rates of 39 percent against 13 percent in the control group. The comparison is stark precisely because the two groups started from the same first stage: similar initial reductions in depression produced durable recovery in one group and a feedback loop in the other, in which treatment-induced assertiveness provoked violence and violence in turn became a source of renewed depressive symptoms.

This finding carries a methodological lesson for the literature on the long-run effects of psychological interventions (Baranov et al., 2020; Blattman et al., 2023). Whether mental health gains persist is not a property of the therapy alone. It depends on how the household responds to the woman's improved wellbeing, and evaluations that stop at symptom scores in the short run will miss the mechanism that determines persistence.

8 Alternative Mechanisms

Our interpretation of the results is that treatment restored women's capacity for self-advocacy, and that the household response to this change depended on the leverage women held. An alternative class of explanations holds that treatment changed the woman in other ways. The group sessions may have built social networks whose value or threat differed by status. Therapy may have shifted women's gender attitudes, so that the conflict we observe reflects a clash of norms rather than voice without leverage. Or treatment may have altered underlying economic preferences, changing bargaining behavior directly. Table 6 tests each of these candidate mechanisms among high schooling women, split by the spousal gap as before.

Table 6: Alternative mechanisms 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: 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 score	-0.03	0.137	517	-0.13	0.149	433	0.342
		(0.120)			(0.135)		
Panel B: Gender Norms							
Gender norms index	0.14	-0.042	161	0.07	-0.137	132	0.658
		(0.145)			(0.129)		
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)		

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.

Panel A examines social capital. The group component of THPP+ brought women into regular contact with peers, and this could affect violence through two channels. Expanded networks could raise the cost of violence by improving women's access to support, which would predict protective effects concentrated where networks grew most. Alternatively, husbands could perceive their wives' new social ties as a threat, generating backlash proportional to social exposure rather than to bargaining position. Neither pattern appears in the data. Treatment effects on women's social interactions and on perceived support from friends are positive but small and statistically insignificant in both groups, and the tests of equality do not reject that the effects are the same ($p = 0.49$ and $p = 0.34$). A channel operating through social capital cannot account for effects on violence that run in opposite directions.

Panel B turns to gender norms. If therapy made women more progressive in their views on women's work, decision-making, or the acceptability of violence, the divergence could reflect differential norm change, with conflict arising where new attitudes collided with old expectations. We find no evidence for this channel. Treatment effects on the gender norms index are small, if anything negative, and statistically indistinguishable from zero in both groups ($\beta = -0.04$ for $W > H$ and $\beta = -0.14$ for $W \leq H$, $p = 0.66$ for their equality). What changed is not what women believed the household should look like but their capacity to act on preferences they already held. This distinction matters for interpretation. The intervention contained no empowerment content, no discussion of women's rights, and no IPV-specific material, and consistent with this design, it moved behavior without moving ideology.

Panel C considers economic preferences. Depression is known to distort risk tolerance and patience, and recovery could shift these primitives, altering labor supply and household negotiation independently of the self-advocacy channel. The estimates on willingness to take risk and to delay gratification are statistically indistinguishable from zero in both groups and from each other, though we note that these outcomes were measured in a single wave with smaller samples, so the estimates are less precise than for the other mechanisms. With that caveat, the absence of detectable preference shifts alongside large behavioral changes is consistent with our reading that treatment restored women's capacity to pursue existing preferences rather than changing the preferences themselves.

Finally, because the intervention was delivered by peer volunteers, its effects could in principle run through who delivered it rather than what it did. The volunteers were local women with an average of 12 years of schooling, most of them married (Appendix Table B19). Figures A7 and A8 examine heterogeneity by peer volunteer characteristics. Treatment effects on depression are stable across volunteer age, marital status, education, motivation, and self-esteem. Effects on violence are also broadly

similar, with one exception: point estimates are larger for women served by volunteers who scored below 75 percent on the post-training competency assessment, though confidence intervals overlap across all comparisons. Two considerations limit what this pattern can explain. Patient characteristics, including baseline violence and the spousal schooling gap, do not predict the type of volunteer a woman was assigned (Appendix Table B20), so volunteer quality cannot generate the heterogeneity by status that organizes our results. And a delivery-quality account offers no reason why the same volunteers would reduce violence for one group of clients while increasing it for another. Taken together, the evidence points to a single conclusion. The intervention worked as intended on the woman. The divergence arose in the household's response to her.

9 Discussion

This paper provides the first experimental evidence on how treating a woman's depression affects her exposure to intimate partner violence. A psychological intervention that durably reduced maternal depression also increased women's exposure to violence on average, and this average conceals two divergent trajectories. Women with more education than their husbands converted their recovery into less sexual violence, greater decision-making power, higher labor force participation, and fertility closer to their stated preferences. Women with equal or lower education experienced escalating violence, retreated from decision-making and work, and ultimately saw their mental health gains unwind. The economic logic is simple. Psychotherapy restores a woman's motivation and capacity to articulate preferences, but it does not by itself change the leverage she brings to the resulting negotiation. Where leverage exists, restored self-advocacy activates it. Where it is absent, restored self-advocacy can result in retaliation. In this framework, mental health and bargaining power are complements — status without the capacity for self-advocacy lies dormant, and self-advocacy without status can backfire.

These results are not an argument against treating depression. The intervention delivered large and lasting mental health benefits, and for a substantial group of women it improved nearly every outcome we measure. The harm arose in a specific configuration, where violence carries little cost for the husband, exit from marriage is not credible, and treatment restored voice without leverage. The risk lies in scaling psychosocial programs blind to this distinction. Our findings suggest pairing treatment with components that shift leverage, such as economic transfers or engagement of husbands and kin, and incorporating violence monitoring into follow-up, since the adverse effects we document emerge with a delay and would be invisible to evaluations that end when

symptom scores improve.

The broader lesson extends beyond mental health. Whether an intervention empowers through psychological recovery, legal rights, or economic transfers, its gains are mediated by the household (Daher et al., 2025). Interventions that change what a woman wants to negotiate, without changing what she can credibly threaten, place her at the mercy of the household's response.

References

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, *100*, 1847–1859. https://doi.org/10.1257/aer.100.4.1847

Anderberg, D., & Rainer, H. (2013). Economic abuse: A theory of intrahousehold sabotage. *Journal of Public Economics*, *97*, 282–295.

Anderberg, D., Rainer, H., Wadsworth, J., & Wilson, T. (2016). Unemployment and domestic violence: Theory and evidence. *The Economic Journal*, *126*, 1947–1979. https://doi.org/10.1111/ecoj.12246

Anderson, S., & Eswaran, M. (2009). What determines female autonomy? evidence from bangladesh. *Journal of Development Economics*, *90*, 179–191. https://doi.org/10.1016/j.jdeveco.2008.10.004

Angelucci, M. (2008). Love on the rocks: Domestic violence and alcohol abuse in rural mexico. *B.E. Journal of Economic Analysis and Policy*, *8*(1). https://doi.org/10.2202/1935-1682.1766

Ashraf, N., Field, E., & Lee, J. (2014). Household bargaining and excess fertility: An experimental study in zambia. *American Economic Review*, *104*(7), 2210–37.

Baranov, V., Bhalotra, S., Biroli, P., & Maselko, J. (2020). Maternal depression, women's empowerment, and parental investment: Evidence from a randomized controlled trial. *American Economic Review*, *110*, 824–859. https://doi.org/10.1257/aer.20180511

Baranov, V., Cameron, L., Contreras Suárez, D., & Thibout, C. (2021). Theoretical underpinnings and meta-analysis of the effects of cash transfers on intimate partner violence in low- and middle-income countries. *The Journal of Development Studies*, *57*(1), 1–25. https://doi.org/10.1080/00220388.2020.1762859

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Beegle, K., Frankenberg, E., & Thomas, D. (2001). Bargaining power within couples and use of prenatal and delivery care in indonesia. *Studies in Family Planning*, *32*, 130–146. https://doi.org/10.1111/j.1728-4465.2001.00130.x

Bhalotra, S., Kambhampati, U., Rawlings, S., & Siddique, Z. (2021). Intimate partner violence: The influence of job opportunities for men and women. *The World Bank Economic Review*, *35*, 461–479. https://doi.org/10.1093/wber/lhz030

Blattman, C., Chaskel, S., Jamison, J. C., & Sheridan, M. (2023). Cognitive behavioral therapy reduces crime and violence over ten years: Experimental evidence. *American Economic Review: Insights*, *5*, 527–545. https://doi.org/10.1257/aeri.20220427

Bloch, F., & Rao, V. (2002). Terror as a bargaining instrument: A case study of dowry violence in rural india. *American Economic Review*, *92*(4), 1029–1043. https://doi.org/10.1257/00028280260344588

Bobonis, G. J., González-Brenes, M., & Castro, R. (2013). Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy*, *5*(1), 179–205. https://doi.org/10.1257/pol.5.1.179

Cullen, C., Alik-Lagrange, A., Ngatia, M., & Vaillant, J. (2025). *The unintended impacts of an intimate partner violence prevention program: Experimental evidence from rwanda*, World Bank Group.

Daher, C. A., Field, E., Swanson, K., & Vyborny, K. (2025). Drivers of change: Employment responses to the lifting of the saudi female driving ban. *American Economic Review*, *115*(9), 3248–3271. https://doi.org/10.1257/aer.20240119

Devries, K. M., Mak, J. Y., Bacchus, L. J., Child, J. C., Falder, G., Petzold, M., Astbury, J., & Watts, C. H. (2013). Intimate partner violence and incident depressive symptoms and suicide attempts: A systematic review of longitudinal studies. *PLoS Medicine*, *10*, e1001439. https://doi.org/10.1371/journal.pmed.1001439

Doepke, M., & Tertilt, M. (2018). Women's empowerment, the gender gap in desired fertility, and fertility outcomes in developing countries. *AEA Papers and Proceedings*, *108*, 358–362. https://doi.org/10.1257/pandp.20181085

Doss, C. (2013). Intrahousehold bargaining and resource allocation in developing countries. *The World Bank Research Observer*, *28*, 52–78. https://doi.org/10.1093/wbro/lkt001

Erten, B., & Keskin, P. (2018). For better or for worse?: Education and the prevalence of domestic violence in turkey. *American Economic Journal: Applied Economics*, *10*, 64–105. https://doi.org/10.1257/app.20160278

Erten, B., & Keskin, P. (2024). Trade-offs? the impact of wto accession on intimate partner violence in cambodia. *Review of Economics and Statistics*, *106*, 322–333. https://doi.org/10.1162/rest_a_01140

Eswaran, M., & Malhotra, N. (2011). Domestic violence and women's autonomy in developing countries: Theory and evidence. *American Economic Review*, *101*, 591–619. https://doi.org/10.1257/aer.101.2.591

Farmer, A., & Tiefenthaler, J. (1997). An economic analysis of domestic violence. *Review of Social Economy*, *55*, 337–358. https://doi.org/10.1080/00346769700000004

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

Guarnieri, E., & Rainer, H. (2021). Colonialism and female empowerment: A two-sided legacy. *Journal of Development Economics*, *151*, 102666. https://doi.org/10.1016/j.jdeveco.2021.102666

Hidrobo, M., Peterman, A., & Heise, L. (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern ecuador. *American Economic Journal: Applied Economics*, *8*(3), 284–303. https://doi.org/10.1257/app.20150048

Jacoby, H. G., & Mansuri, G. (2010). Watta satta: Bride exchange and women's welfare in rural pakistan. *American Economic Review*, *100*, 1804–1825. https://doi.org/10.1257/aer.100.4.1804

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: Relation to cognitive inhibition. *Cognition & Emotion*, *24*, 281–298. https://doi.org/10.1080/02699930903407948

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

Leykin, Y., Roberts, C. S., & DeRubeis, R. J. (2011). Decision-making and depressive symptomatology. *Cognitive Therapy and Research*, *35*, 333–341. https://doi.org/10.1007/s10608-010-9308-0

Li, Z. (2026). What if she educates more? examining the dynamics of intrahousehold bargaining power [Early view. Check final volume and pages before submission.]. *Review of Development Economics*. https://doi.org/10.1111/rode.70161

Lundberg, S., & Pollak, R. A. (1993). Separate spheres bargaining and the marriage market. *Journal of Political Economy*, *101*, 988–1010. https://doi.org/10.1086/261912

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

Nezu, A. M. (1986). Negative life stress and anxiety: Problem solving as a moderator variable. *Psychological Reports*, *58*, 279–283. https://doi.org/10.2466/pr0.1986.58.1.279

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

Oram, S., Fisher, H. L., Minnis, H., Seedat, S., Walby, S., Hegarty, K., Rouf, K., Angénieux, C., Callard, F., Chandra, P. S., et al. (2022). The lancet psychiatry commission on intimate partner violence and mental health: Advancing mental health services, research, and policy. *The Lancet Psychiatry*, *9*(6), 487–524.

Rapp, D., Zoch, B., Khan, M. M. H., Pollmann, T., & Krämer, A. (2012). Association between gap in spousal education and domestic violence in india and bangladesh. *BMC Public Health*, *12*, 467. https://doi.org/10.1186/1471-2458-12-467

Rasul, I. (2008). Household bargaining over fertility: Theory and evidence from malaysia. *Journal of Development Economics*, *86*, 215–241. https://doi.org/10.1016/j.jdeveco.2007.02.008

Roy, S., Hidrobo, M., Hoddinott, J., & Ahmed, A. (2019). Transfers, behavior change communication, and intimate partner violence: Postprogram evidence from rural bangladesh. *Review of Economics and Statistics*, *101*(5), 865–877. https://doi.org/10.1162/rest_a_00791

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

Spencer, C., Mallory, A., Toews, M., Stith, S., & Wood, L. (2022). Mental health factors and intimate partner violence perpetration and victimization: A meta-analysis. *Trauma, Violence, & Abuse*, *23*(4), 809–824. https://doi.org/10.1177/1524838019882364

Stevenson, B., & Wolfers, J. (2006). Bargaining in the shadow of the law: Divorce laws and family distress. *The Quarterly Journal of Economics*, *121*, 267–288. https://doi.org/10.1093/qje/121.1.267

Tauchen, H. V., Witte, A. D., & Long, S. K. (1991). Domestic violence: A nonrandom affair. *Demography*, *28*(3), 453–475. https://doi.org/10.2307/2006669

Tol, W. A., Murray, S., Lund, J. K., Bass, J. K., Bolton, J., Murray, L., Jordans, M. J. D., et al. (2019). Can mental health treatments help prevent or reduce intimate partner violence in low- and middle-income countries? a systematic review. *The Lancet Psychiatry*, *6*(10), 869–881. https://doi.org/10.1016/S2215-0366(19)30180-2

Trevillion, K., Oram, S., Feder, G., & Howard, L. M. (2012). Experiences of domestic violence and mental disorders: A systematic review and meta-analysis. *PLoS ONE*, *7*, e51740. https://doi.org/10.1371/journal.pone.0051740

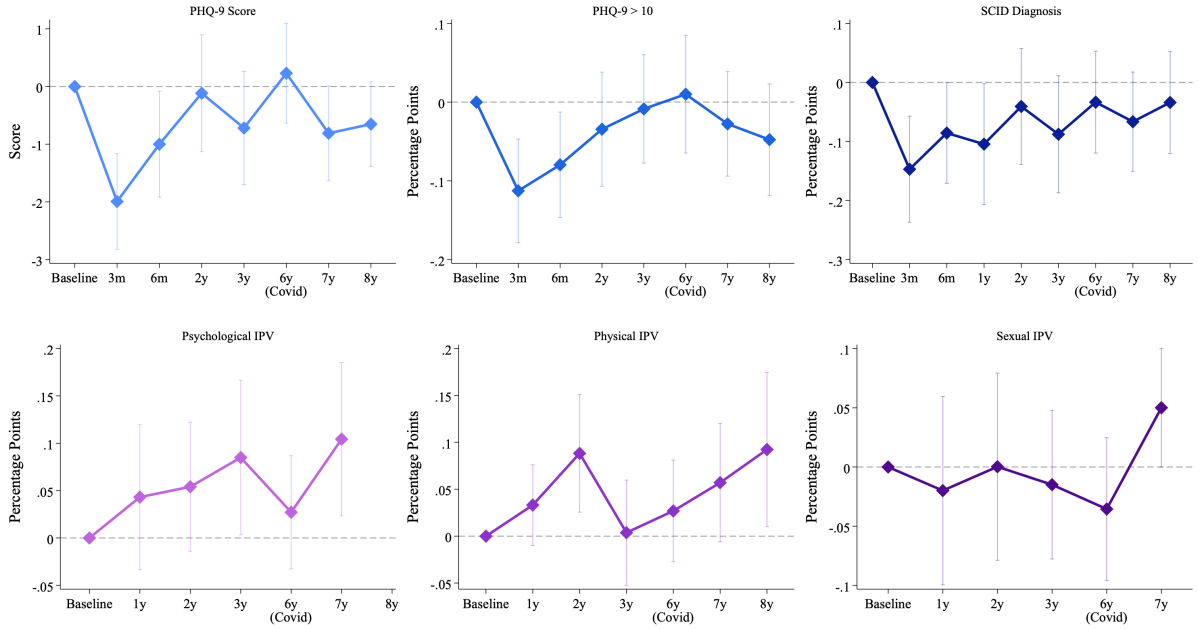
Turner, E. L., Sikander, S., Bangash, O., Zaidi, A., Bates, L., Gallis, J., Ganga, N., O'Donnell, K., Rahman, A., & Maselko, J. (2016). The effectiveness of the peer-delivered thinking healthy plus (thpp+) program for maternal depression and child socioemotional development in pakistan: Study protocol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

Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., et al. (2012). Years lived with disability (ylds) for 1160 sequelae of 289 diseases and injuries 1990–2010: A systematic analysis for the global burden of disease study 2010. *The Lancet*, *380*(9859), 2197–2223. https://doi.org/10.1016/S0140-6736(12)61729-2

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

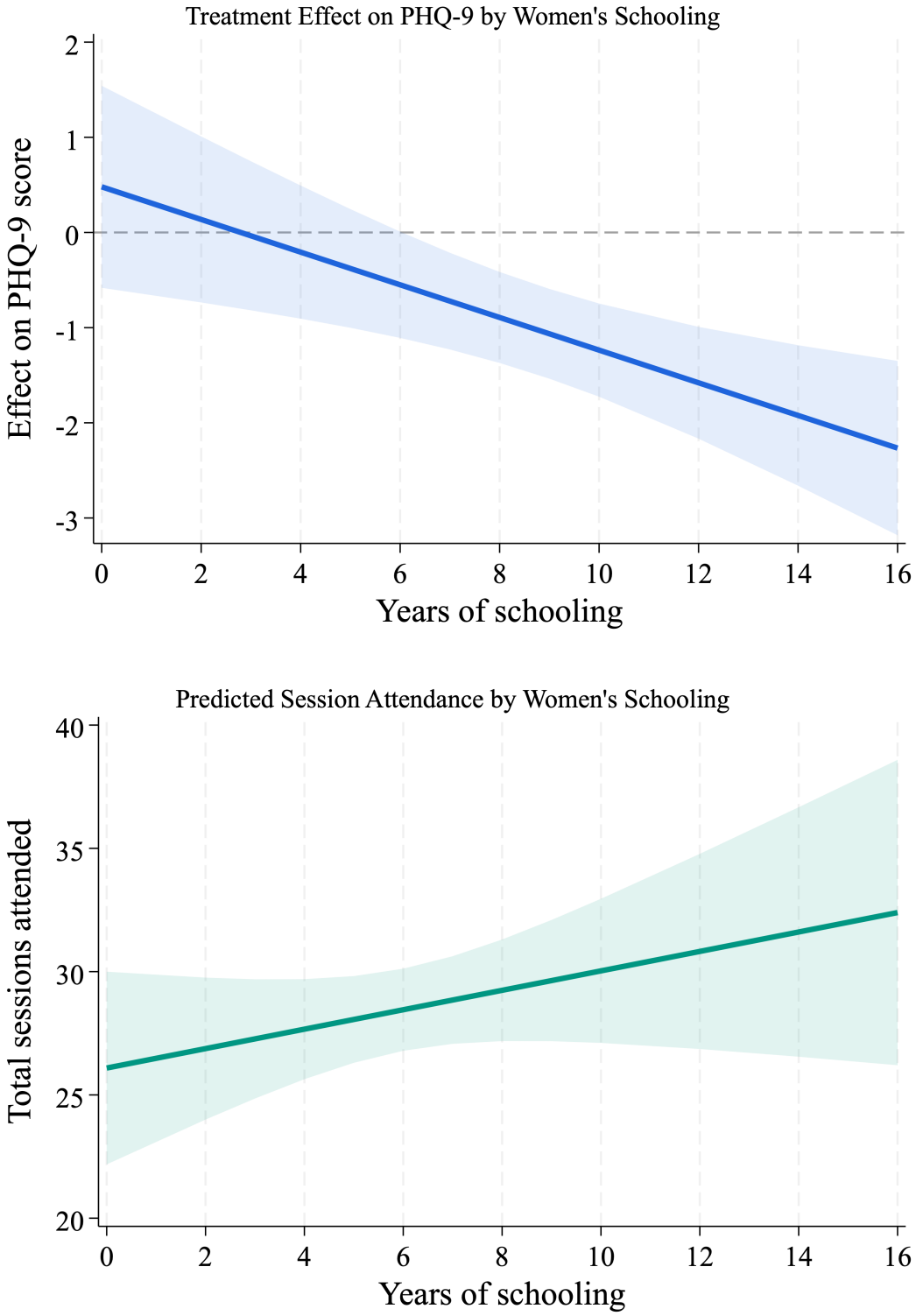
Appendix A - Additional Figures

Figure A1: Trajectories of Depression and IPV



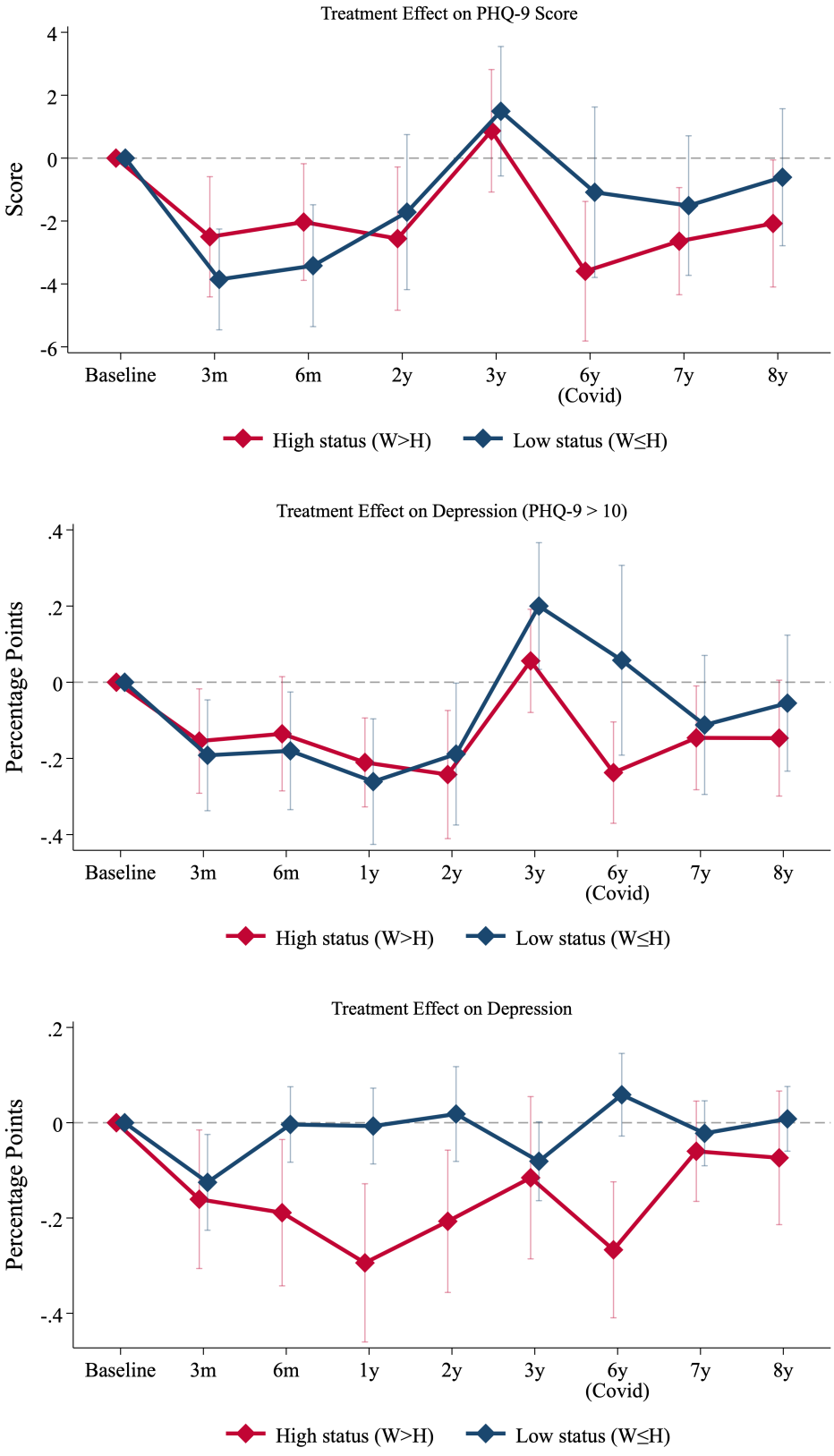
Notes: The above figures plot differences in trajectories between treatment and control arms for depression and the probability of experiencing IPV over time.

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



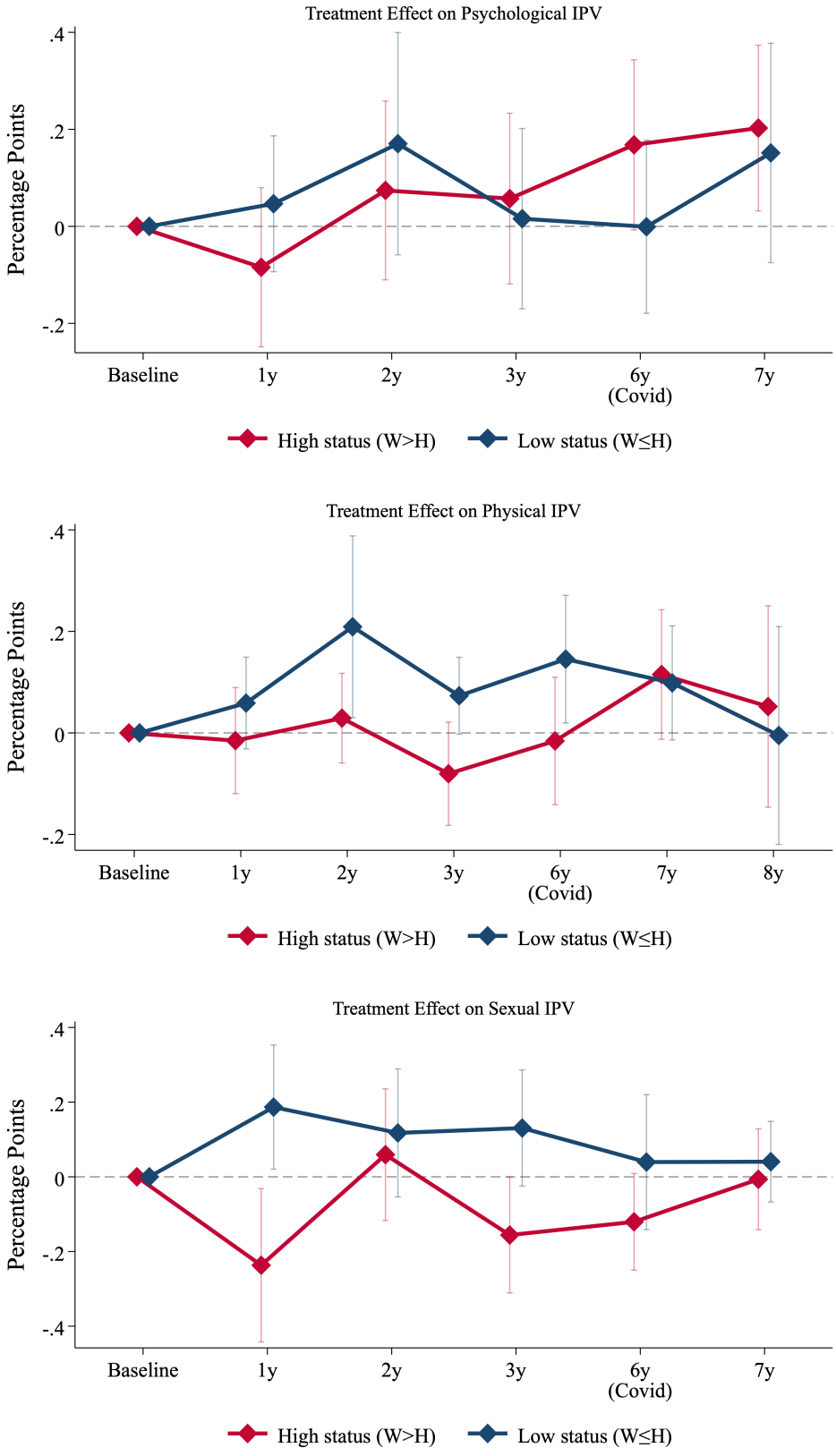
Notes: The top figure plots marginal treatment effects of CBT on PHQ-9 scores by women's absolute schooling levels with 90% CI. The bottom figure tests the correlation between the number of psychotherapy sessions attended and women's schooling levels.

Figure A3: Trajectories of depression by spousal gap - high schooling women



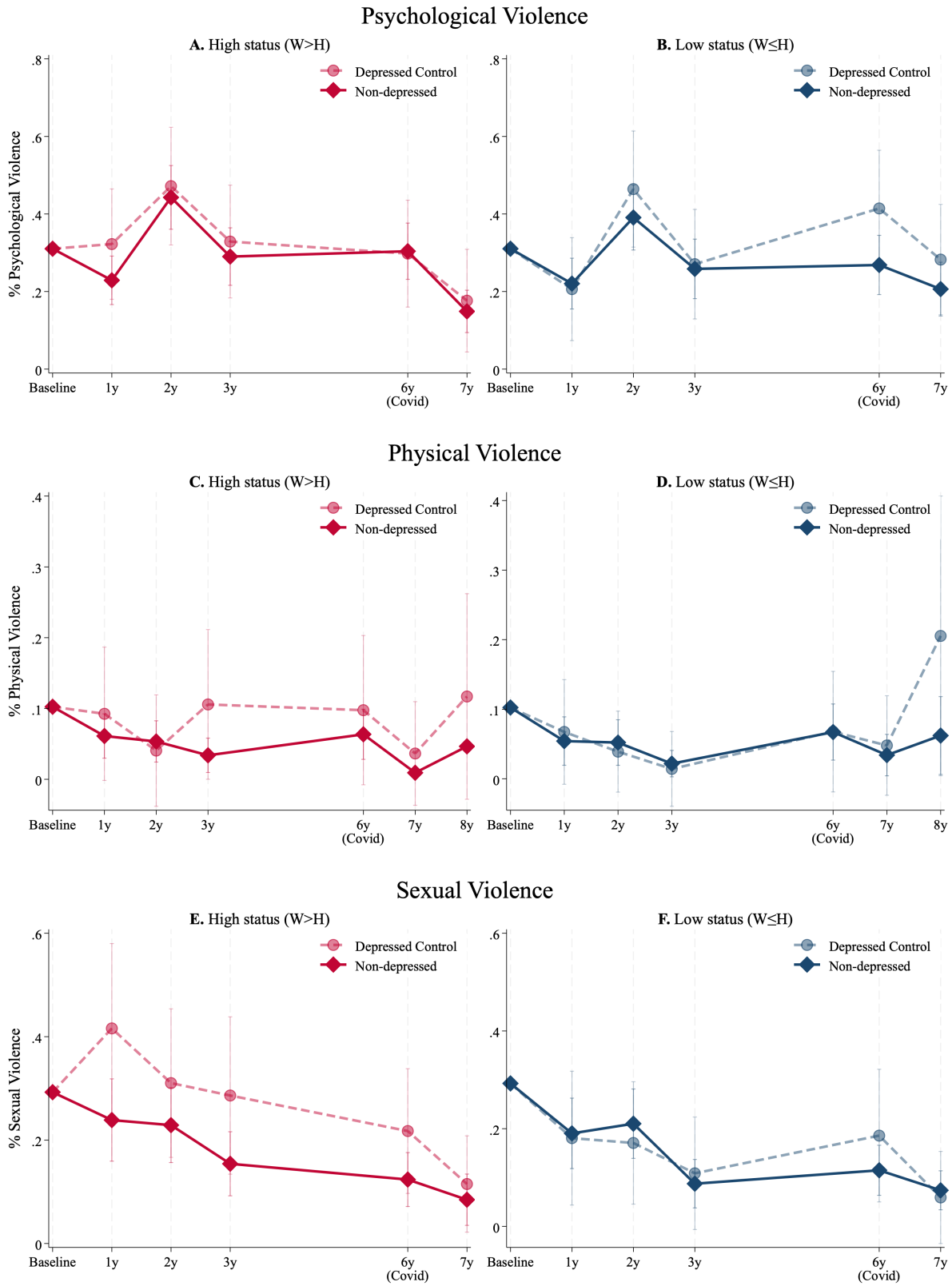
Notes: The above figures plot the treatment effects (difference between treatment and control) on mental health, physical and sexual IPV by spousal schooling gap among high absolute schooling women only.

Figure A4: Trajectories of IPV by spousal gap - high schooling women



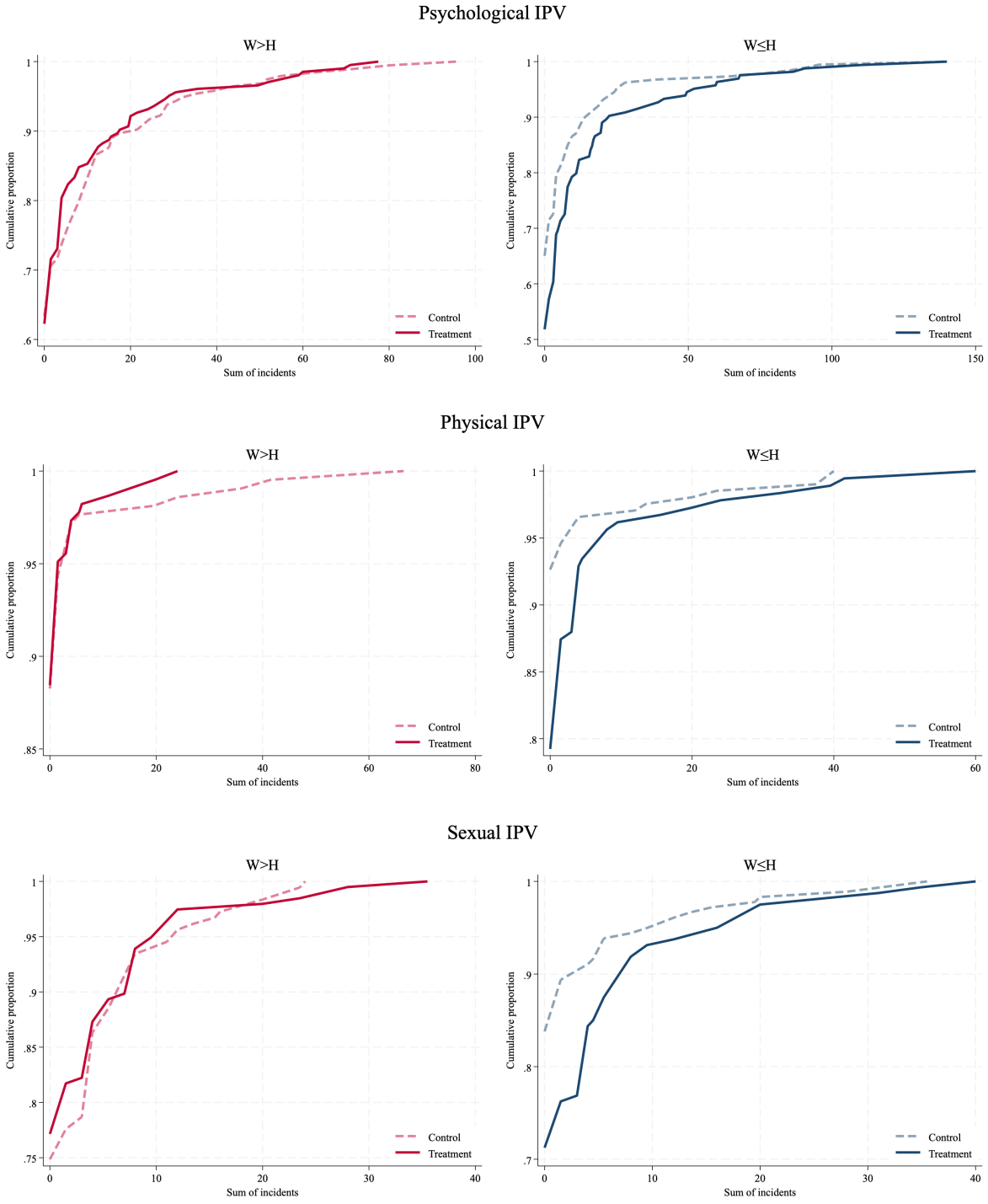
Notes: The above figures plot the treatment effects (difference between treatment and control) on mental health, physical and sexual IPV by spousal schooling gap among high absolute schooling women only.

Figure A5: Trajectories of IPV by spousal gap - non depressed and depressed control



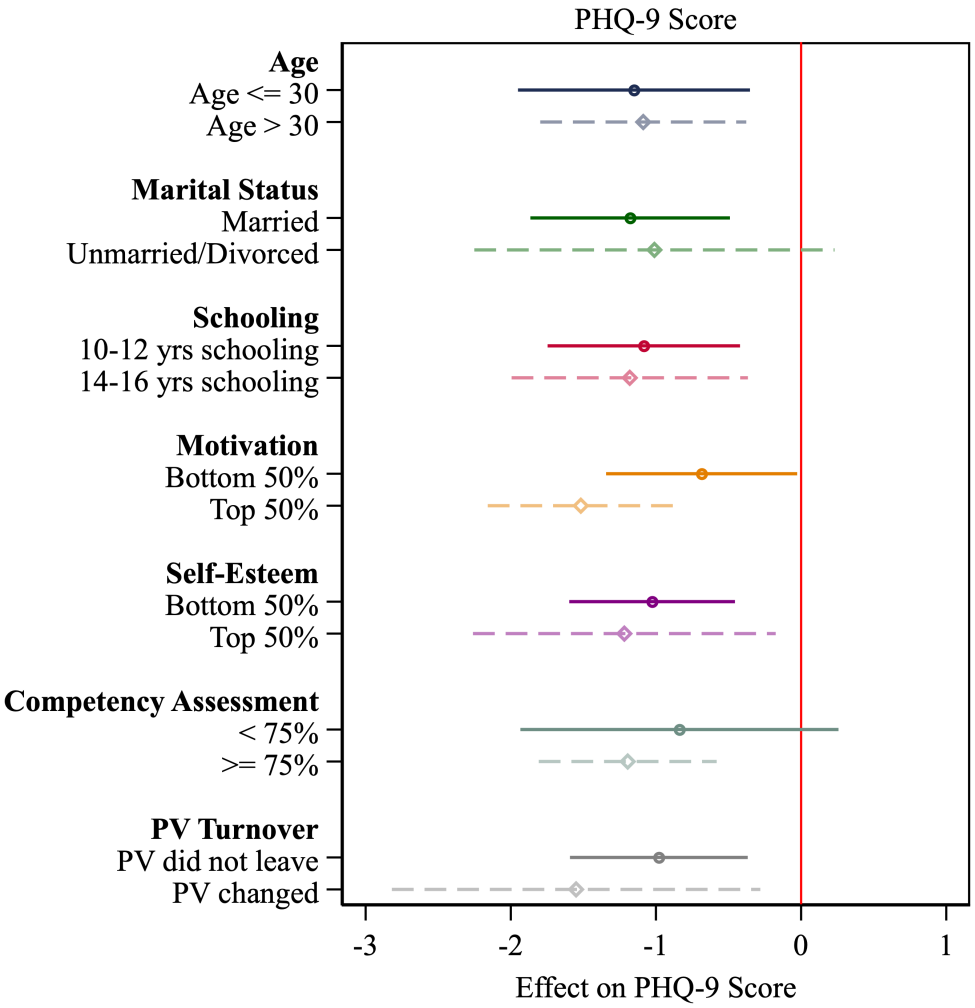
Notes: The above figures plot the raw trajectories of IPV: proportion of women experiencing psychological, physical and sexual in the sample by women's status relative to their husbands, restricted to women with high absolute school and comparing the non-depressed cohort to the depressed control arm.

Figure A6: Frequency of IPV by arm and spousal gap - high schooling women



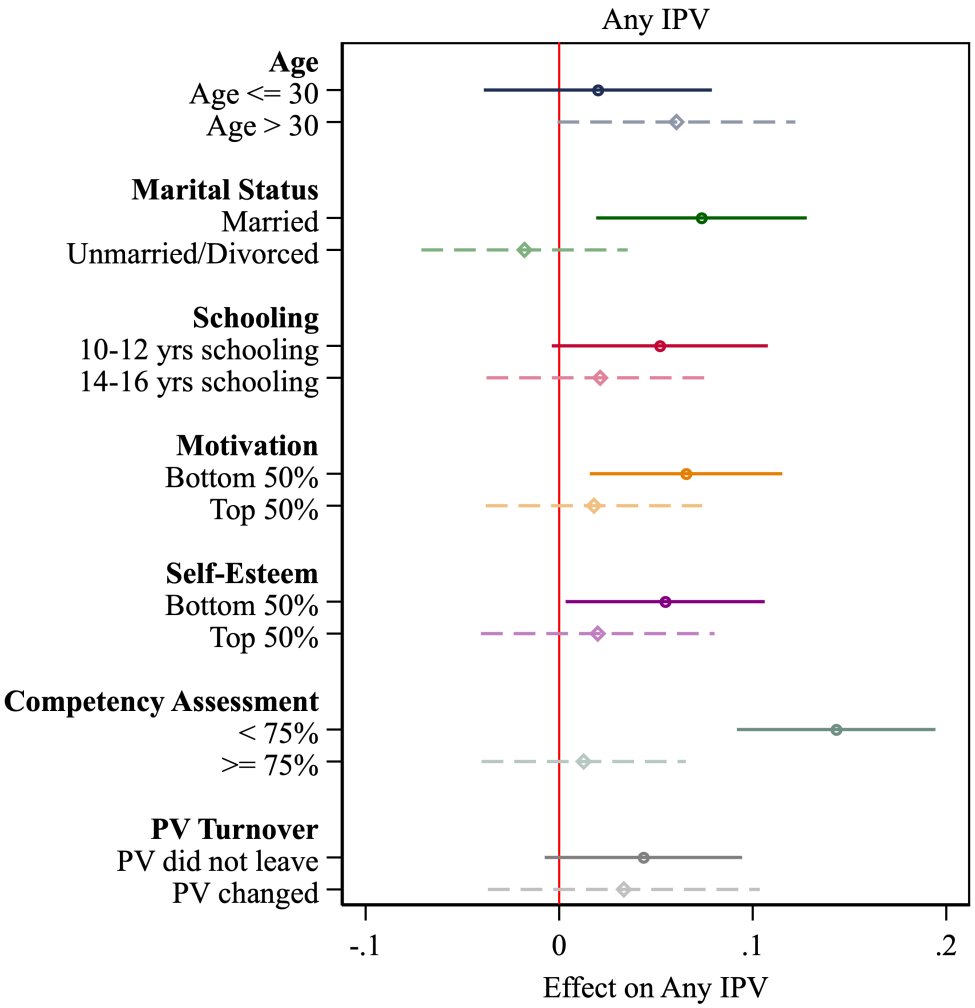
Notes: The above figures plot the raw trajectories of IPV: proportion of women experiencing psychological, physical and sexual in the sample by treatment arm and women's status relative to their husbands over the eight years of follow-up, restricted to women with high school or more education.

Figure A7: Treatment effects on depression scores by peer volunteer characteristics



Notes: The above figures plot the coefficient estimates for the effects of treatment on depression scores through various types of peer volunteers.

Figure A8: Treatment effects on violence by peer volunteer characteristics



Notes: The above figures plot the coefficient estimates for the effects of treatment on the probability of experiencing any type of intimate partner violence through various types of peer volunteers.

Appendix B – Additional Tables

B1 Baseline balance by women's relative and absolute schooling 45
B2 Balance in baseline characteristics across follow-up waves - Full Sample 46
B3 Balance in baseline characteristics across follow-up waves - High schooling $W > H$ 47
B4 Balance in baseline characteristics across follow-up waves - High schooling $W \leq H$ 48
B5 Balance in attritor characteristics - Full Sample 49
B6 Balance in attritor characteristics - High absolute schooling and $W > H$. 50
B7 Balance in attritor characteristics - High absolute schooling and $W \leq H$. 51
B8 Main outcome measures used in the analyses 52
B9 Average treatment effects on mental health over time 53
B10 Average treatment effects on violence over time 54
B11 Indirect versus direct endorsement of physical violence 55
B12 Baseline characteristics by relative and absolute schooling 56
B13 Post baseline characteristics of women in the control arm 57
B14 Correlation between spousal schooling gap and female agency 58
B15 Fertility since index child by schooling gap 59
B16 Violence effects by women's absolute and relative schooling 60
B17 WHO IPV Items by schooling gap - High schooling women only 61
B18 Effects on relationship quality - High schooling women only 62
B19 Characteristics of Peer Volunteers 63
B20 Do patient characteristics predict PV characteristics? 64

Appendix Table B1: Baseline balance by women's relative and absolute schooling

	High Schooling				Low Schooling			
	W > H N=120		W ≤ H N=97		W > H N=50		W ≤ H N=303	
	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 B2: Balance in baseline characteristics across follow-up waves - Full Sample

	6 Mths N=458		Year 1 N=451		Year 2 N=430		Year 3 N=422		Year 6 N=432		Year 7 N=403		Year 8 N=400	
	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	-0.46	0.32	-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	-1.50	0.22	-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	-2.16	0.03	-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	-0.56	0.65	-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	0.26	0.46	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	0.68	0.42	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	1.06	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	0.26	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	0.69	-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	0.36	-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	0.79	-0.02	0.60	-0.01	0.78
Total adults in the household	-0.33	0.22	-0.33	0.22	-0.22	0.42	-0.29	0.29	-0.11	0.70	-0.25	0.38	-0.26	0.36
First child	-0.00	0.96	0.01	0.90	-0.01	0.76	-0.02	0.58	-0.00	0.95	-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	0.58	-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	0.06	0.54	0.06	0.60	0.05	0.62
People per room	0.32	0.04	0.34	0.03	0.33	0.05	0.31	0.06	0.34	0.04	0.36	0.04	0.30	0.08
SES asset index	-0.13	0.40	-0.13	0.42	-0.16	0.33	-0.17	0.30	-0.10	0.53	-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	0.21	0.62	-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	-0.65	0.05	-0.82	0.02	-0.82	0.02
W > schooling than H	0.07	0.10	0.07	0.09	0.06	0.21	0.06	0.17	0.07	0.13	0.07	0.12	0.07	0.12
Any psychological IPV past 12m	0.02	0.60	0.00	0.95	0.02	0.65	0.01	0.77	-0.01	0.90	-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	0.04	0.30	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	-0.08	0.98	-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	1.02	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	2.86	0.02	3.20	0.01	3.72	0.00
Joint test (*p*-value)		0.17		0.17		0.09		0.22		0.20		0.03		0.00

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

Appendix Table B3: Balance in baseline characteristics across follow-up waves - High schooling W > H

	6 Mths N=96		Year 1 N=92		Year 2 N=87		Year 3 N=85		Year 6 N=84		Year 7 N=80		Year 8 N=81	
	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	-0.67	0.46	-0.52	0.57	-0.80	0.37	-0.40	0.66	-0.30	0.75	-0.45	0.64	-0.42	0.66
Mother's height (cm)	-0.47	0.70	-0.18	0.88	0.17	0.90	-0.13	0.92	0.26	0.84	0.67	0.61	0.53	0.68
Mother's weight (kg)	-3.95	0.13	-3.67	0.18	-4.25	0.13	-3.51	0.21	-2.93	0.30	-2.14	0.45	-2.28	0.42
Mother's waist circ. (cm)	-3.45	0.10	-2.69	0.21	-3.71	0.10	-2.66	0.24	-2.08	0.36	-1.77	0.44	-1.89	0.40
Mother's blood pressure	-1.51	0.54	-1.29	0.61	-1.72	0.51	-1.74	0.51	-1.45	0.60	-1.70	0.54	-1.06	0.70
PHQ-9 Score	-0.05	0.93	-0.06	0.93	-0.29	0.65	-0.42	0.56	-0.23	0.74	-0.48	0.51	-0.53	0.45
WHODAS Score	-0.06	0.97	0.06	0.98	0.06	0.98	-0.34	0.86	-0.07	0.97	-0.38	0.85	-0.37	0.85
PSS Score	-2.66	0.08	-2.50	0.11	-1.95	0.22	-1.56	0.34	-2.00	0.23	-2.00	0.23	-2.30	0.17
Current major dep. episode	0.13	0.16	0.09	0.33	0.11	0.28	0.05	0.59	0.07	0.45	0.05	0.63	0.06	0.58
Nuclear family	0.12	0.08	0.10	0.17	0.12	0.12	0.11	0.15	0.08	0.29	0.10	0.22	0.10	0.23
Joint/extended family	-0.18	0.06	-0.16	0.10	-0.19	0.05	-0.18	0.07	-0.16	0.11	-0.20	0.05	-0.19	0.06
Grandmother present	-0.13	0.11	-0.11	0.20	-0.14	0.10	-0.11	0.20	-0.10	0.26	-0.10	0.27	-0.09	0.29
Total adults in the household	-0.48	0.41	-0.60	0.31	-0.32	0.60	-0.54	0.38	-0.07	0.91	-0.50	0.42	-0.37	0.55
First child	-0.01	0.96	-0.00	0.99	-0.04	0.73	-0.03	0.76	-0.01	0.90	-0.03	0.82	-0.01	0.93
Number of boys	0.03	0.85	0.03	0.87	0.04	0.84	0.06	0.76	0.02	0.93	0.00	1.00	-0.01	0.94
Number of girls	0.05	0.79	0.10	0.61	0.15	0.41	0.08	0.66	0.12	0.54	0.17	0.37	0.16	0.41
People per room	-0.11	0.65	-0.09	0.71	-0.07	0.79	-0.12	0.64	0.07	0.78	0.10	0.68	0.12	0.63
SES asset index	0.28	0.32	0.30	0.31	0.48	0.09	0.46	0.11	0.52	0.07	0.39	0.17	0.42	0.14
Mother's education	-0.80	0.07	-0.84	0.06	-0.62	0.17	-0.68	0.15	-0.74	0.11	-1.02	0.03	-1.01	0.03
Father's education	-0.94	0.10	-0.82	0.15	-0.97	0.10	-0.98	0.10	-0.87	0.14	-1.00	0.11	-0.95	0.12
Any psychological IPV past 12m	-0.07	0.53	-0.10	0.37	-0.03	0.81	-0.02	0.83	-0.06	0.60	-0.09	0.43	-0.10	0.38
Any physical IPV past 12m	-0.09	0.26	-0.10	0.25	-0.01	0.87	-0.05	0.60	-0.04	0.67	-0.07	0.45	-0.07	0.42
Any sex-related IPV past 12m	0.19	0.05	0.18	0.06	0.26	0.01	0.18	0.08	0.23	0.03	0.17	0.11	0.19	0.08
Frequency of psychological IPV	-8.99	0.05	-9.54	0.04	-6.28	0.19	-8.31	0.10	-7.59	0.13	-8.85	0.10	-9.00	0.09
Frequency of physical IPV	-0.46	0.45	-0.49	0.45	-0.14	0.83	-0.23	0.74	-0.27	0.70	-0.25	0.73	-0.28	0.70
Frequency of sex-related IPV	3.27	0.15	3.15	0.19	5.10	0.04	4.50	0.07	4.62	0.07	4.83	0.07	4.67	0.07
Joint test (*p*-value)		0.04		0.10		0.02		0.12		0.12		0.16		0.13

Note: This table tests for balance in baseline characteristics between arms across follow-up waves exclusively for women with high levels of absolute schooling and W>H (greater schooling relative to husband).

Appendix Table B4: Balance in baseline characteristics across follow-up waves - High schooling W $\leq$ H

	6 Mths N=80		Year 1 N=78		Year 2 N=74		Year 3 N=72		Year 6 N=72		Year 7 N=67		Year 8 N=65	
	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	-0.48	0.60	-0.17	0.85	-0.38	0.69	0.23	0.80	-0.25	0.79	-0.30	0.76	-0.21	0.83
Mother's height (cm)	1.09	0.39	1.43	0.27	1.08	0.43	1.34	0.33	1.43	0.30	1.74	0.23	2.30	0.12
Mother's weight (kg)	3.54	0.21	4.33	0.13	4.38	0.14	4.79	0.10	5.37	0.06	5.70	0.06	5.70	0.07
Mother's waist circ. (cm)	2.45	0.27	2.91	0.21	3.11	0.19	3.45	0.15	3.65	0.11	3.92	0.11	3.70	0.14
Mother's blood pressure	4.36	0.19	3.29	0.33	4.63	0.20	4.12	0.26	4.78	0.19	4.42	0.25	5.57	0.16
PHQ-9 Score	1.74	0.05	1.54	0.08	1.52	0.10	1.67	0.07	1.77	0.05	1.47	0.12	1.30	0.17
WHODAS Score	1.56	0.46	2.00	0.34	2.11	0.31	2.36	0.26	2.30	0.29	3.11	0.17	3.29	0.15
PSS Score	2.49	0.15	2.32	0.18	2.07	0.23	2.17	0.22	2.34	0.18	2.14	0.25	2.10	0.26
Current major dep. episode	0.19	0.05	0.18	0.07	0.20	0.04	0.27	0.01	0.25	0.01	0.22	0.04	0.24	0.03
Nuclear family	-0.18	0.03	-0.21	0.02	-0.23	0.01	-0.22	0.02	-0.26	0.01	-0.29	0.00	-0.28	0.00
Joint/extended family	0.09	0.38	0.15	0.15	0.17	0.10	0.15	0.18	0.16	0.13	0.18	0.12	0.17	0.13
Grandmother present	-0.04	0.70	-0.00	0.98	0.01	0.91	-0.02	0.85	0.02	0.82	0.03	0.79	0.02	0.87
Total adults in the household	0.27	0.64	0.29	0.62	0.38	0.53	0.15	0.79	0.13	0.83	0.17	0.78	-0.13	0.83
First child	0.04	0.68	0.05	0.61	0.01	0.90	-0.05	0.56	0.02	0.82	0.01	0.91	0.00	1.00
Number of boys	0.06	0.76	0.07	0.72	0.12	0.56	0.15	0.47	-0.00	0.99	-0.12	0.56	-0.11	0.59
Number of girls	-0.02	0.93	-0.04	0.86	-0.00	0.98	0.08	0.74	0.05	0.84	0.07	0.77	0.11	0.65
People per room	0.22	0.40	0.16	0.54	0.07	0.79	0.16	0.57	0.10	0.74	0.18	0.55	0.14	0.66
SES asset index	-0.46	0.13	-0.23	0.48	-0.29	0.39	-0.23	0.47	-0.34	0.31	-0.17	0.61	-0.07	0.83
Mother's education	-0.19	0.39	-0.21	0.33	-0.18	0.43	-0.28	0.23	-0.29	0.22	-0.43	0.07	-0.45	0.07
Father's education	0.05	0.91	0.23	0.58	0.24	0.58	0.31	0.49	0.03	0.95	-0.02	0.97	-0.07	0.87
Any psychological IPV past 12m	0.27	0.02	0.26	0.02	0.22	0.06	0.23	0.05	0.22	0.07	0.19	0.13	0.16	0.22
Any physical IPV past 12m	0.07	0.43	0.12	0.16	0.10	0.26	0.10	0.29	0.11	0.23	0.11	0.27	0.10	0.32
Any sex-related IPV past 12m	0.03	0.78	0.06	0.61	0.06	0.61	0.08	0.49	0.06	0.62	0.08	0.55	0.05	0.72
Frequency of psychological IPV	14.78	0.05	17.26	0.02	15.10	0.04	16.78	0.03	15.56	0.04	16.35	0.05	12.95	0.11
Frequency of physical IPV	2.83	0.08	3.21	0.05	3.14	0.05	3.09	0.07	3.03	0.08	3.00	0.10	2.95	0.12
Frequency of sex-related IPV	1.05	0.71	1.84	0.52	1.01	0.72	1.56	0.60	1.19	0.69	1.72	0.59	1.94	0.54
Joint test (*p*-value)		0.09		0.08		0.15		0.05		0.04		0.06		0.07

Note: This table tests for balance in baseline characteristics between arms across follow-up waves exclusively for women with high levels of absolute schooling and W$\leq$H (less or equal schooling relative to husband).

Appendix Table B5: Balance in attritor characteristics - Full Sample

	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 B6: Balance in attritor characteristics - High absolute schooling and W > H

	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	0.87	0.76	1.86	0.47	1.33	0.56	2.03	0.35	1.99	0.34	1.79	0.35	1.82	0.36
Mother's height (cm)	1.96	0.51	2.67	0.34	2.45	0.26	1.65	0.44	2.66	0.24	3.04	0.15	2.88	0.18
Mother's weight (kg)	-2.76	0.64	-0.47	0.93	-3.09	0.48	-0.45	0.92	1.60	0.72	2.70	0.54	2.54	0.57
Mother's waist circ. (cm)	-0.80	0.88	2.11	0.66	-1.90	0.63	1.30	0.74	3.12	0.42	3.25	0.40	3.09	0.43
Mother's blood pressure	0.80	0.91	0.90	0.89	0.71	0.90	0.64	0.90	1.24	0.80	0.97	0.84	2.16	0.64
PHQ-9 Score	-0.91	0.56	-0.67	0.61	-0.70	0.59	-1.14	0.28	-0.77	0.50	-1.02	0.33	-1.21	0.25
WHODAS Score	-5.58	0.19	-3.82	0.23	-2.57	0.43	-3.49	0.22	-2.98	0.31	-2.93	0.28	-3.12	0.26
PSS Score	-2.09	0.54	-0.47	0.87	0.64	0.82	1.72	0.50	0.82	0.75	0.83	0.74	0.21	0.93
Current major dep. episode	-0.09	0.59	-0.20	0.21	-0.09	0.49	-0.24	0.09	-0.20	0.16	-0.22	0.09	-0.22	0.10
Nuclear family	-0.24	0.22	-0.25	0.13	-0.14	0.33	-0.16	0.25	-0.25	0.08	-0.18	0.16	-0.19	0.14
Joint/extended family	0.33	0.12	0.31	0.10	0.17	0.35	0.21	0.23	0.27	0.11	0.16	0.31	0.18	0.25
Grandmother present	0.33	0.12	0.31	0.10	0.11	0.53	0.21	0.23	0.27	0.11	0.21	0.17	0.24	0.14
Total adults in the household	2.56	0.03	1.67	0.14	1.80	0.06	1.28	0.17	2.47	0.01	1.25	0.19	1.56	0.09
First child	0.11	0.58	0.05	0.81	-0.07	0.69	-0.04	0.83	0.02	0.93	-0.01	0.94	0.02	0.91
Number of boys	-0.60	0.06	-0.47	0.11	-0.37	0.16	-0.31	0.19	-0.40	0.09	-0.39	0.08	-0.43	0.06
Number of girls	-0.02	0.94	0.13	0.63	0.24	0.44	0.04	0.90	0.10	0.75	0.21	0.44	0.18	0.51
People per room	-0.00	1.00	0.05	0.91	0.18	0.66	0.03	0.95	0.46	0.23	0.49	0.22	0.52	0.20
SES asset index	0.93	0.06	0.75	0.09	1.00	0.04	0.90	0.06	1.01	0.03	0.58	0.23	0.66	0.17
Mother's education	-0.69	0.52	-0.49	0.61	0.40	0.65	0.20	0.80	-0.01	0.99	-0.39	0.60	-0.43	0.57
Father's education	0.07	0.96	0.74	0.60	0.03	0.98	0.13	0.90	0.47	0.68	0.19	0.85	0.28	0.79
Any psychological IPV past 12m	-0.09	0.69	-0.10	0.61	0.09	0.62	0.08	0.64	0.01	0.96	-0.06	0.71	-0.08	0.62
Any physical IPV past 12m	-0.31	0.10	-0.25	0.13	0.04	0.78	-0.05	0.74	-0.02	0.86	-0.08	0.54	-0.09	0.49
Any sex-related IPV past 12m	0.07	0.76	-0.04	0.83	0.20	0.26	-0.03	0.85	0.07	0.69	-0.06	0.71	-0.03	0.84
Frequency of psychological IPV in the past 12 months	-1.10	0.09	-0.89	0.11	-0.09	0.86	-0.37	0.40	-0.28	0.56	-0.34	0.39	-0.37	0.36
Frequency of physical IPV in the past 12 months	-0.31	0.22	-0.24	0.26	-0.03	0.86	-0.06	0.71	-0.08	0.62	-0.06	0.70	-0.07	0.66
Frequency of sexual IPV in the past 12 months	0.68	0.28	0.44	0.41	1.15	0.02	0.78	0.10	0.79	0.09	0.69	0.11	0.68	0.12
Joint test (*p*-value)		0.17		0.32		0.02		0.10		0.07		0.11		0.00

Note: This table tests for differential attrition across follow-up waves by treatment arm for women with high absolute schooling and high relative schooling as compared to the husband.

Appendix Table B7: Balance in attritor characteristics - High absolute schooling and W $\leq$ H

	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	0.09	0.96	0.32	0.87	-0.24	0.86	0.96	0.52	-0.14	0.92	-0.57	0.66	-0.04	0.97
Mother's height (cm)	-8.33	0.00	-6.61	0.00	-6.70	0.00	-5.93	0.00	-5.47	0.00	-4.32	0.02	-2.94	0.10
Mother's weight (kg)	-2.95	0.51	-2.26	0.64	-2.16	0.61	-2.56	0.56	-0.22	0.96	-0.87	0.83	-0.72	0.85
Mother's waist circ. (cm)	3.58	0.45	1.73	0.69	1.97	0.62	1.35	0.72	2.63	0.54	1.92	0.61	1.45	0.68
Mother's blood pressure	6.33	0.26	0.73	0.90	4.37	0.29	2.03	0.65	4.24	0.27	1.93	0.62	3.80	0.28
PHQ-9 Score	-0.11	0.95	-1.98	0.27	-1.80	0.22	-1.54	0.29	-1.28	0.38	-2.18	0.13	-2.19	0.13
WHODAS Score	-2.76	0.48	-0.69	0.85	-0.49	0.90	-0.95	0.80	-0.51	0.88	0.11	0.97	1.09	0.73
PSS Score	3.20	0.29	0.70	0.83	-0.62	0.86	-0.83	0.79	-0.06	0.99	-0.87	0.76	-0.84	0.76
Current major dep. episode	0.15	0.51	0.09	0.62	0.11	0.56	0.27	0.13	0.22	0.17	0.08	0.60	0.10	0.48
Nuclear family	0.27	0.18	0.11	0.31	0.08	0.35	0.09	0.27	0.01	0.96	-0.04	0.72	0.02	0.90
Joint/extended family	-0.20	0.43	0.09	0.62	0.11	0.56	0.03	0.85	0.06	0.70	0.08	0.60	0.04	0.82
Grandmother present	0.05	0.87	0.07	0.78	0.11	0.57	0.01	0.95	0.13	0.49	0.12	0.49	0.09	0.61
Total adults in the household	-0.91	0.53	-0.72	0.55	-0.45	0.66	-0.81	0.48	-1.01	0.37	-0.61	0.56	-1.29	0.20
First child	-0.32	0.19	-0.18	0.43	-0.29	0.15	-0.39	0.05	-0.21	0.29	-0.18	0.31	-0.21	0.21
Number of boys	0.05	0.91	-0.03	0.93	0.11	0.79	0.16	0.66	-0.27	0.45	-0.48	0.17	-0.40	0.25
Number of girls	0.41	0.42	0.20	0.67	0.29	0.47	0.39	0.30	0.38	0.37	0.35	0.38	0.41	0.27
People per room	-0.41	0.57	-0.47	0.43	-0.77	0.18	-0.36	0.51	-0.61	0.20	-0.35	0.41	-0.44	0.27
SES asset index	-0.81	0.27	0.10	0.87	-0.04	0.95	0.08	0.89	-0.19	0.71	0.17	0.76	0.41	0.43
Mother's education	0.20	0.77	0.16	0.78	0.22	0.65	0.03	0.95	-0.04	0.92	-0.20	0.62	-0.25	0.50
Father's education	-1.11	0.27	-0.07	0.94	-0.11	0.88	0.13	0.86	-0.71	0.32	-0.57	0.43	-0.71	0.29
Any psychological IPV past 12m	0.05	0.87	0.00	1.00	-0.18	0.42	-0.17	0.43	-0.20	0.34	-0.26	0.16	-0.31	0.09
Any physical IPV past 12m	0.18	0.30	0.33	0.06	0.18	0.29	0.13	0.37	0.17	0.26	0.11	0.43	0.08	0.54
Any sex-related IPV past 12m	-0.39	0.13	-0.22	0.37	-0.18	0.42	-0.13	0.54	-0.19	0.36	-0.16	0.38	-0.18	0.31
Frequency of psychological IPV in the past 12 months	0.04	0.96	0.50	0.46	-0.19	0.80	-0.06	0.93	-0.22	0.73	-0.30	0.60	-0.64	0.29
Frequency of physical IPV in the past 12 months	0.24	0.44	0.31	0.26	0.17	0.64	0.07	0.82	0.05	0.87	-0.04	0.88	-0.07	0.78
Frequency of sexual IPV in the past 12 months	-0.42	0.35	-0.06	0.88	-0.54	0.42	-0.38	0.55	-0.49	0.43	-0.39	0.46	-0.20	0.71
Joint test (*p*-value)		0.29		0.87		0.40		0.49		0.39		0.40		0.00

Note: This table tests for differential attrition across follow-up waves by treatment arm for women with high absolute schooling and less or equal schooling relative to the husband.

Appendix Table B8: Main outcome 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 B9: Average 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 B10: Average 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 B11: Indirect versus direct endorsement of physical violence

	Overall (1)	Intervention (2)	Control (3)
List experiment (indirect)	0.07	0.28	0.07
Direct question	0.07	0.16	0.09
Difference (indirect − direct)	0.00	0.12	-0.02
Observations	403	194	209

Notes: This table reports prevalence estimates of physical IPV (slap) from a list randomization experiment conducted at the six-year follow-up. Respondents were randomly assigned to receive a list of either four or five statements; the additional statement concerned whether the respondent had been slapped by her husband. The list experiment estimate is the difference in mean endorsement counts between the five-item and four-item groups. The full analysis is available in [collaborators' paper, in preparation].

Appendix Table B12: Baseline characteristics by relative and absolute schooling

	High Schooling		Low Schooling	
	W > H	W ≤ H	W > H	W ≤ H
	(1)	(2)	(3)	(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 B13: 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 B14: 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 B15: 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 B16: 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 B17: 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 B18: 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 index	-0.31	0.324	0.03	-0.439**	0.031
		(0.272)		(0.188)	
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 situations	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 B19: 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 B20: 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.