Participatory Theater Empowers Women: Evidence from West Bengal, India*

Saraswata Chaudhuri† Karla Hoff‡ Jyotsna Jalan§ Sattwik Santra¶

June 8, 2025

Abstract

Domestic violence is socially and economically costly, yet common and widely accepted in many countries. This paper evaluates an intervention by *Jana Sanskriti* (*JS*) - a non-governmental organization based in rural West Bengal, a state in eastern India that uses community-based participatory theater to reduce violence against women. Using a primary survey of 3,000+ married women between the ages 18-49 years and their husbands in 92 villages, we find, on average, *JS* reduces the proportion of married women who experience physical abuse by 9 percentage points from a baseline of 33 percent. Outcomes such as sexual abuse, knowledge of laws against domestic abuse, willingness to report abuse, and joint household decision making are also "positively" affected by *JS*. Since the *JS* intervention is not a randomized controlled trial, we obtain these causal effects by adjusting for potential confounders, that *JS* informed us about, between the intervention and potential outcomes. Different sensitivity analyses suggest that only very strong unadjusted/unobserved confounders (not communicated to us by *JS*) can overturn the causal results reported in our paper. Such strong confounders are implausible since if *JS* had not informed us of a confounder, then that is unlikely to be strong. In other words, our causal results are robust to plausible inadequacies of the observed confounders in making the intervention "as if random".

JEL Classification: C52, J12, J16, K42, L82, Z13

Keywords: domestic violence, narratives, endogenous treatment, potential outcomes, gender, patriarchy, sensitivity analysis

*The World Bank Research Support Budget and the Centre for Training and Research in Public Finance and Policy, Kolkata provided financial support. We benefitted from comments by Isaiah Andrews, Sheheryar Banuri, Kaushik Basu, Sourav Bhattacharya, Annette Brown, Stefan Dercon, Victoire Girard, Willemien Kets, Michal Kolesar, Sandip Mitra, Dilip Mookherjee, Victor Orozco, Joan Ricart-Huguet, Arijit Sen, Moses Shayo, Rohini Somanathan, Petra Todd, Ralph Yarrow, Victoria Zinde-Walsh among others. This work would not have been possible without the support of Jana Sanskriti, our field investigators, and Jharna Panda, who supervised the survey with great diligence.

†Department of Economics, McGill University & Cireq, Montreal. Email: saraswata.chaudhuri@mcgill.ca.

‡Department of Economics, Columbia University. Email: kh2830@columbia.edu.

§Centre for Studies in Social Sciences, and Centre for Training and Research in Public Finance and Policy, Kolkata. Email: jjalan@cssscal.org.

¶Centre for Training and Research in Public Finance and Policy, Kolkata. Email: sattwik@cssscal.org.

1 Introduction

Man is an animal suspended in webs of significance he himself has spun.

– Clifford Geertz, *The Interpretation of Cultures*

Domestic violence is a global problem that contributes to women's and children's death and illnesses and to emotional disorders that are associated with the perpetuation of violence by the children in their adulthood (Egeland, 1993). Yet domestic violence is both prevalent and widely accepted (Doepke et al., 2012). Among women in low-income and lower middle-income countries, the median proportion who believe that a husband is justified to hit or beat his wife under some circumstances (such as arguing with him) is 58 and 33 percent respectively (Table 1). In many societies, men risk a loss of honor if they do not beat a recalcitrant wife (Derne, 1994). A common expression of men in North India is "What kind of a man is he who does not hit?" (Chowdhry, 2015). A statement by an Ethiopian woman illustrates attitudes widely shared in rural areas: "It is sometimes necessary for husbands to beat their wives when they commit mistakes to correct them…it is also a sign of strong manhood" (Narayan et al., 2000).

Change in social acceptance of domestic violence requires a change in social norms. It requires that people come to see women as having basic entitlements as individuals with rights to make decisions about their lives. But if sufficient members of a community accept a prescriptive norm that denies women such rights, then even a community member who does not share that belief will have an incentive to follow the norm to avoid social sanctions. Such sanctions are common. Therefore, changing social norms require changing the opinions and biases of the society at large instead of a single person. One can start with the individual or the community, but it is crucial to influence *both*. Under international pressure, many countries enacted, for the first time in the 1990's and 2000's, laws against domestic violence (Htun and Jensenius, 2020). Yet, the social acceptance of domestic violence undermines the enforcement of laws designed to safeguard the well-being of its victims.

Evaluations of legal prohibitions of domestic violence, information campaigns and messages against domestic violence embedded in entertainment in mass media find that such interventions have had little or no effect in reducing domestic violence ((Banerjee et al., 2019a) and (Cooper et al., 2020)). They are unlikely to change the behavior, if they do not trigger community-wide discussions that change the prescriptive norms. Furthermore, in the context of intimate partner violence there are two sets of people with opposing and conflicting views (Agüero, 2019). In this case providing simple information about the rights and protections under the law for one agent often implies characterizing widely accepted gender norms as wrong. To do that in isolation could lead to more domestic violence between partners.

Forum Theater is a potential tool that can change community norms and beliefs. The Brazilian political activist and drama theorist Augusto Boal (1979) developed Forum Theater as a tool in the Theater of the Oppressed. He was influenced by Paulo Freire (1970), who argued that to learn how to stop oppression, an individual must be a *co-creator* of the knowledge and develop the self-assertiveness to act on it. Forum Theater provides a space for dialogue between actors and the audience. It tries to induce men and women, little by little, to think critically about their prescribed roles and try out forms of rebellious action onstage, in front of the audience and actors who have been trained to respond in ways that reflect widely shared village beliefs. Boal viewed this as a rehearsal for real social change.

Today there are several successful applications of Forum Theatre all over the world over a range of issues - cultural, political, and social differences and demands. It has been used by peasants in the villages of South 24 Parganas in West Bengal to address the pros and cons of traditional and modern agricultural practices in their play *Unnoyoner Pothe* (in translation, "On the Path to Development"), by workers in *Amra Jekhane Dariye* (in translation, "Where We Stand") to address the "unholy" nexus between trade union leaders and industrialists in the organized sector with the workers having no decision-making power, among others. Participatory theatre has been used in other countries too to address social issues. For

example, in *Woarabeba*, a fishing community in Winneba, Ghana, received environmental health education through participatory theatre. The local community did benefit from such theatre because people started using the toilets instead of defecating outside (Asante and Zakaria, 2021). In another example, Forum Theater was used in Malaysia to impart moral education to high school children (Thambu et al., 2020).

Despite the wide use of Forum Theater under different scenarios, no rigorous and large-scale evaluation of its impact on reforming social issues has been done (Scharinger, 2013). This paper seeks to fill that gap by evaluating Forum Theater, designed to overcome the limitations of non-participatory interventions to change individuals' beliefs and induce broad socio-economic change.

We evaluate the impact of Jana Sanskriti (*JS*, in translation, "People's Culture") a non-governmental organization (NGO) in West Bengal, India. *JS* is one of the largest organizations in the world that performs Forum Theater. It is a non-profit, independent network of community-based theater troupes. *JS* started performing Forum Theater in the late 1990's and has grown organically in West Bengal by drawing on interested individuals as actors and as organizers of troupes. In 2015, there were 500-600 actors in 13 troupes with a roughly equal gender balance. Each of the troupes performed regularly in 12-19 villages (Yarrow, 2017). The actors resided in the areas where their troupes performed. Most of the actors were from families of agricultural or wage laborers.

The troupes perform stark dramas based on conflicts that most people have heard about or experienced themselves, and *JS* dramatists have learnt about in private and/or group meetings. In repeat performances in a village, a play might be adapted to take account of intervening actions within the village or with the authorities. The plays give individuals an opportunity to learn to think critically about oppressive practices that occur in their own villages and, in some cases, to respond to events as they develop in real time. Supplemental Appendix B presents synopsis of three plays that are typical of the dramas that *JS* dramatists write.

In this paper, we evaluate the impact of village exposure to *JS* plays on the empowerment of women and on attitudes towards gender roles in the villages. *JS* has performed Forum Theater in villages in the Sundarban delta in the South 24 Parganas district of West Bengal since 1998 (see Figure 1). By the time of our survey in 2016, *JS* troupes were performing in about 125 villages. Once *JS* had started performing in a village, it continued to perform there regularly. Thus, selection of the villages exposed to *JS* was a one-time decision.

We randomly selected 31 villages in which *JS* had performed and 61 villages in which *JS* had never performed. We call the first set of villages "treatment villages" and the second set of villages "control villages". In each village in our sample, we use the voters list for the 2011 legislative assembly election in West Bengal to randomly select 3000+ households. In the sampled households, we administered in 2016, a survey in private, to the wife (between 18-49 years) and to her husband. The outcomes that we use are village-level averages of the individual responses. The women and their husbands were asked the same questions on issues concerning women's autonomy, except for those related to actual incidents of abuse. Only the wives were asked these questions to reduce the risk of repercussions on them for participating in a survey that covered topics related to domestic violence.

The econometric strategy adopted in this paper to evaluate the impact of *JS'* intervention is based on the Neyman-Rubin potential outcomes framework. We define the treatment effect as the difference in the potential outcome(s), like, say, physical abuse, with and without the treatment, i.e., with and without exposure of the village to *JS*' performances. It is of course not possible to observe both potential outcomes, i.e., with and without treatment, for any individual village. Hence, as is standard, we focus on identifying and estimating the causal parameters that are the common aggregate/average measures of the treatment effect.

The *JS* intervention is however not a randomized controlled trial (RCT) as *JS* did not randomly select the villages where it performs. *JS* explained to us that they sought to perform in relatively populous villages in South 24 Parganas that were geographically distant from towns, and not well connected to the outside world. It is possible that such villages

had more conservative beliefs about gender roles. Hence, there could be causal paths to the potential outcomes from these pretreatment variables that *JS* informed us that they had considered while selecting the villages where they would perform. These variables could be potential confounders and, thus, we adjust/control for them to circumvent potential selection bias in identifying the causal parameters under a selection on observables assumption.

Our estimates suggest for example, that on average, exposure of a village to *JS*' performances reduced the proportion of village households where wives reported being physically abused by their husbands by about 9.3 percentage points from an estimated baseline mean of 32.6 percentage points (mean potential outcome without treatment). This reduction is about 80 percent of the standard deviation of the observed physical abuse. The average effect was a little larger on the treated villages. Based on the wive's and husbands' responses we find similar effectiveness of the exposure of a village to *JS*'s performances in reducing not only other forms of domestic abuse but also various indicators of patriarchal norms like limited or no participation in making decisions related to the visiting wife's family or relatives, number of children to have, use of contraceptives among others.[1]

Point-identification of these causal parameters is based on our selection on observables assumption that all the confounders between the treatment and the potential outcomes are adjusted for. This assumption cannot be validated statistically. However, this assumption is violated if there persists confounders unadjusted/uncontrolled for in our analysis. Therefore, we use sensitivity tools recently proposed by Masten and Poirier (2018, 2020) and Chernozhukov et al. (2024) to evaluate whether our results are robust to violations of the selection on observables assumption. We report the magnitude of association that would be needed between possible unadjusted/uncontrolled confounders with the treatment and with the potential outcomes to cause enough bias in our causal estimates to overturn their

[1]We obtain very similar estimates based on the estimators like regression adjustment, inverse probability weighting, combination of regression and inverse probability weighting, matching, and also double-debiased machine learning estimators based on covariate selections using lasso, that are commonly used in empirical analysis. Our discussion will focus on the regression adjustment estimator for concreteness, but for completeness we will also report in Supplemental Appendix A the results for some of these other estimators.

signs, and lets the reader decide the plausibility of such strength of association. Given what we know about *JS*' treatment assignment mechanism, we argue by using "benchmarking" exercises that such strength of association with the unadjusted/uncontrolled confounder – often referred to as "unobserved confounder" henceforth – would have to be unrealistically strong to overturn our causal results for many of the outcomes of interest.

Our paper is organized as follows. Section 2 gives background information on *JS* and on the impact of narratives and other mechanisms by which community-based Forum Theater influences behavior and norms. Section 3 explains our sampling and econometric strategies. Section 4 presents the estimated causal effects. Section 5 presents sensitivity analysis of those effects. Section 6 concludes. The Appendix reports the tables and figures discussed in the paper. Supplemental Appendices A, B, C, and D report, respectively, further estimation and sensitivity results, synopses of *JS* plays, our survey design, and our survey questionnaire.

2 Background

2.1 *Jana Sanskriti*

In public performances, *JS* dramatizes personal struggles faced by the villagers. Supplemental Appendix B presents synopsis of three plays by *JS* that illustrate the stark dramas that *JS* presents. In preparation for writing a play on a particular topic, *JS* holds discussion groups in villages to which anyone can come and recount personal problems on that topic. Then in theater workshops, *JS* members debate the social context of a problem, such as domestic violence. The artistic director arranges themes and vignettes depicted by *JS* members into a narrative. He refrains from offering any final solution to the problem, for it is the audience, with the help of a facilitator called the Joker who asks them questions, to rescript the play so that the oppression stops (Mohan, 2010). The goal is to lead villagers to imagine household and village relations differently at a particular point in the performed narrative when the options for change in the outcome are not foreclosed.

Many *JS* plays are a response to specific local issues; the plays bring widely known events directly into the performance for elaboration and discussion. For instance, *JS* performed a play on rape in response to an incident where neighbors, despite knowing that a rape was occurring, took no action (Ganguly, 2017). A play is generally repeated in the same location within 1-3 months (Yarrow, 2017). In repeat performances, it may be adapted to take account of intervening action within the village or with the authorities. That can make the atmosphere energized, since some people may be considering what to do in an ongoing situation.

A *JS* performance is preceded by music, dance, and games to create a playful and trusting atmosphere. That is followed by an uninterrupted enactment of a play depicting a conflict that is very familiar in the village. After the uninterrupted performance, the Joker will ask whether everyone in the audience agrees with the actions that were taken onstage. Some people will usually say no. The Joker will then explain that the play will begin again and any member of the audience can, at any point, interrupt the performance by shouting "Stop," and go onstage. At this point, the volunteer puts on the costume of the character that they want to play, and indicates the moment in the play from which they want the scene to be taken forward. (If no one volunteers to go onstage, then at some point the Joker will shout "Stop" to interrupt the performance. The actors will then freeze in positions that express the conflictual situation of the drama until someone from the audience volunteers to go onstage.) In a single performance of participatory theater, a series of volunteers will go onstage. They will try out strategies to avert the oppression (Boal, 1979). Because the volunteers are both spectators of the play and actors in the scene that they interrupt, Boal (1979) coined the term spect-actors to describe them.

Participation by members of the audience bftext - both as actors onstage and as commentators on the performance - makes a performance both a play and its analysis. As spect-actors, members of the audience have the opportunity to try to persuade oppressors to change their behavior. Watching the spect-actors and actors rescript the play, other mem-

bers of the audience may question, for the first time, norms that they have long taken for granted. The uninterrupted part of the performance may be as short as 20 minutes. The interactive part of the performance can last 2-3 hours (Ganguly, 2010). A series of spect-actors will add new dimensions to the situation, and actors and the audience will step out of the pretense of the story to discuss possible turns of the plot.

2.2 Mechanisms in Forum Theater to foster social change

Forum Theater uses three principal mechanisms to induce social change: reframing of issues, information transmission, and the rehearsal of social change. We discuss each, in turn.

Social reframing: Forum Theater triggers discussion between the actors and the audience about social norms. Spect-actors put themselves in the shoes of a character and try out new ways to behave. From their actions and words, together with everyone else's reactions and words, a collectively created new ending to the story emerges. The Joker plays a key role in bringing individuals' comments to some resolution, and ensuring that individuals respond to each other.

In the past decade, economists for the first time have recognized that narratives provide mental frames through which people interpret reality. By reframing situations, new stories can have a large impact on behavior and aggregate outcomes. Benabou et al. (2018) show how new stories that raise self-image concerns influence the social norms that individuals follow. Stories influence macroeconomic fluctuations (Shiller, 2019), work effort (Akerlof and Rayo, 2020), health choices (Banerjee et al. (2018); Banerjee et al. (2019a)), fertility rates (La Ferrara et al., 2012), the ability to perform (Hoff and Stiglitz, 2010) and racial hate and discrimination (Ang (2023); Esposito et al. (2023)). Akerlof (2020) characterizes the neglect of stories in mainstream economics as a "sin of omission." The unique role of Forum Theater is to *give members of a community the opportunity to collectively change stories that justify their subjugation*. Forum Theater makes possible collective story editing. It redistributes to oppressed groups an important power that is normally held only by socially dominant

groups—the power of cultural representation.

Transmitting information and eliciting empathy: Ideology can block the absorption of information that would correct false beliefs on which the ideology is based (Kahan et al., 2017). Exposure and experience can pierce the ideological blinders. In a field experiment in Israel, Jha and Shayo (2019) show that a few weeks of actively trading Israeli and Palestinian financial assets changed the participants' beliefs about the sociotropic costs of war, which shifted the participants' votes in the 2015 Israeli election toward parties more supportive of the peace process. Engagement in the stories of *JS* leads some in the audience to see for the first time the suffering of victims of domestic violence and feel empathy for them. It can improve decision making by letting people think about themselves in the third person and, thus, probably with greater objectivity. After watching a play on domestic violence, one member of the audience encountered the leading actress at a bus stop and pledged to her, "Didi (sister), I will not beat my wife again. I beat her quite often. When you were crying after being beaten by your husband in the play, I remembered my wife. She cries exactly like that when I beat her." (p. 30, Ganguly (2010)).

"Performing agency": After a behavior has become routinized and automatic, a higher level of engagement than passive viewing or listening is necessary to induce behavior change: individuals must not only think differently, they must also practice acting differently (Bourdieu, 1977). Forum Theater gives people a relatively safe space in which to discuss and analyze their responses and those of others and to intervene. The Joker is required not to comment on the proceedings but to ask questions to advance the discussion and to ensure fairness so that all voices are heard. Individuals rehearse social change onstage, which (Yarrow, 2012) describes as "performing agency." Evidence of intense parasocial engagement is that a member of the audience once threw a brick at an actor who was about to sell a child into slavery (Mills, 2009).

The audience learns immediately some of the effects of new strategies from the responses of the actors. *Spect-actors* may get social validation from their interactions with the actors

and the applause of the audience. This may bolster the courage to adopt new behaviors. It is especially important for women to have the experience of "performing agency" or seeing their peers do it, since in rural India women are trained from childhood not to argue or be assertive.

3 Sampling and Econometric Strategies

3.1 Sampling strategy for our survey

From the late-1990s to date, *JS* has regularly performed in about 125 villages in the South 24 Parganas district of West Bengal (Figure 1). These villages are located in the Sundarbans delta. As noted above, we call these villages the population of our treatment villages.

The choice of the "controls" is a fundamental issue in the impact evaluation of any treatment. One should choose the control villages to be as similar as possible to the treatment villages in terms of pretreatment characteristics such as remoteness, population, communications facilities, etc. that, as far as we know, influenced *JS*' treatment assignment mechanism. The goal is to make the treatment assignment "as if random" once these characteristics are adjusted for. One could then choose as controls the nearby villages in the same three administrative blocks of the treatment villages. That is, however, not ideal for an intervention like the *JS* performances for which spillover effects are natural. Moreover, that might not isolate non-exposed from exposed villages and enforce compliance since it might be easy to travel from the control to the treatment villages to witness *JS*' performance. To avoid spillovers and noncompliance, we chose the control villages from the three neighboring administrative blocks since travel between blocks is generally difficult due to the presence of water bodies. The villages in these three neighboring blocks are our population control villages.

We drew a random sample of 31 villages from the population of treatment villages where *JS* had performed for at least 10 years as of the time of our survey. We drew a sample of 61 villages from the population of control villages. To select the sampled villages, we used

stratified random sampling described in Supplemental Appendix C. As shown in Figure 1, there are no contiguous control and treatment villages in our sample. About 90 percent of the sampled respondents in the treatment villages had heard of *JS*. No respondent in the control villages had heard of *JS*. This makes large spillovers of the impact of a village's exposure to *JS* unlikely. To the extent that such spillovers occur, our estimates of the impacts of village exposure to *JS* are biased toward zero.

Our analysis for impact evaluation focuses on the combined population of the treatment and control villages. This population is representative of the relatively more populous (in 1991) villages in the Sundarbans delta in the South 24 Parganas district of West Bengal.[2]

To estimate *JS*' impact on abuse and on its social acceptance, we need village-level indicators that are not available from official sources. Between March 2014 and March 2016, we administered a survey across selected villages. Supplemental Appendix D presents the survey questionnaire. Our sample consists of 1,635 married couples from 31 treatment villages, and 1,846 married couples from 61 control villages. The average age of female (male) respondents was 31 (37) in the treatment villages and 30 (36) in the control villages. Table 2 presents the basic descriptive statistics of the relevant variables — the outcomes of our interest (from our survey), and the covariates that likely influenced the treatment assignment by *JS* (from Census 1991) — in the treatment and control villages.

3.2 Empirical framework

Let $D_i \in \{0, 1\}$ be the treatment indicator, i.e., $D_i = 1$ if *JS* has been performing in village i since 1998, and $D_i = 0$ if *JS* has never performed in village i. *JS* never stopped performing in a village once it had started performing there. To fix ideas, we continue the exposition with one of the outcomes – "physical abuse". This variable is defined as the proportion of households in a village where a married woman reported that she had been physically abused

[2]To match closely with the treatment villages, we chose control villages to be also more populous than villages on average in the Sundarbans delta in the South 24 Parganas. The control villages happen to be on average even more populous and much more dense (see Table 2 for density) than the treatment villages.

by her husband. Our discussion applies similarly to the other outcomes of interest.

Following the Neyman-Rubin framework, let $Y_i(1)$ and $Y_i(0)$ denote the potential "physical abuse" in village i with and without treatment, respectively. In our context, it makes sense to think about the potential outcomes for both treatment and control villages given their geographical proximity and similarity in social, economic and demographic conditions. We assume no spillovers, i.e., the potential outcomes of any village are assumed not to depend on another village's exposure/nonexposure to the *JS* plays.

For each village i, we observe: $Y_i = (1 - D_i)Y_i(0) + D_iY_i(1) = Y_i(0) + D_i(Y_i(1) - Y_i(0))$. Since we can never observe $Y_i(1)-Y_i(0)$, i.e., the effect for any village i, our causal parameters of interest are the standard aggregate measures of the effect of *JS*: (i) the average effect: ATE $= E[Y_i(1)-Y_i(0)]$, (ii) the average effect on the treated: ATT $= E[Y_i(1)-Y_i(0)|D_i = 1]$, and the average effect on the untreated: ATUT $= E[Y_i(1) - Y_i(0)|D_i = 0]$.

These aggregate causal parameters are not necessarily equal since the *JS* intervention was not an RCT. We focus the discussion in the text in terms of ATE since it can better inform policies for broader implementation of forum theatre to reduce domestic abuse and shift patriarchal norms (Table 1). We also report the ATT and ATUT for completeness since their identification and estimation do not require additional assumptions.[3] We do not however emphasize ATT/ATUT because, unlike in studies where ATT is of substantive interest since their treatment allocation are based on pretreatment outcomes or the expected future outcomes, the choice of treatment villages in our study is based on *JS*'s idiosyncratic criteria – populous, not well-connected, remote villages – that may not have broader appeal.

More commonly and noncontroversially, one could interpret our reported ATE (or ATT/ATUT) as the effects in our sample and not in any population (Imbens, 2004); our standard errors are then conservative, i.e., the effects are in fact more significant than reported.

[3]To remain conservative about the significance of the estimated effects, we will not use finite-population correction of the standard errors by multiplying them with the less-than-unit factor $\sqrt{(N-n)/(N-1)}$, where N is the total number of villages in this finite population and n is the number of sampled villages.

3.3 Identification and Estimation

The central empirical challenge is that *JS* did not randomly choose the villages where it performs. We discussed several times with the leaders of *JS* whether there were any systematic criteria for selecting the villages where *JS* regularly performs. *JS*' response was that they sought to perform in relatively populous villages that were geographically distant from towns, and not well connected to the outside world. These characteristics could affect the potential outcomes $Y_i(0)$ and $Y_i(1)$ of the *JS* intervention. With that in mind, we use data from the Census of India (1991) on (log) population density of the village, (log) distance to the nearest town, access to communication facilities (e.g., post, telegraph and telephone facility), fraction of girls among children under 6 years, and literacy growth rate between 1981 and 1991 as potential confounders between the potential outcomes and *JS*'s treatment assignment, i.e., selection of villages to perform plays since 1998.

Denoting these observed confounders collectively as W_i for village i, to point-identify ATE/ATT/ATUT, we make the assumption of the selection on observables (conditional independence/ ignorability/ unconfoundedness; see p. 54, Angrist and Pischke (2009)):

$$Y_i(0) \text{ and } Y_i(1) \text{ are independent of } D_i, \text{ conditional on } W_i \text{ for all } i. \tag{1}$$

These W_i's are pretreatment variables affecting the choice of treatment villages in 1998 and hence are not "bad" controls in the sense of Angrist and Pischke (2009) (Section 3.2.3). Also, we cannot rule out direct causal paths from these W_i's to the potential outcomes of interest, and hence these W_i's are not instrumental variables satisfying "exclusion restrictions".[4] In other words, these W_i's are strong candidates for "good" controls (confounders) that must be adjusted for to circumvent the problem of omitted variable bias. Such adjustment/balancing/controlling seems necessary because while the distributions of the W_i's in

[4]Furthermore, controlling for these pretreatment variables will most likely not cause the so-called M-bias since it seems perhaps difficult to conceptualize a plausible causal path between the treatment and the outcomes via the pretreatment variables (W_i) as "colliders" on the path (Cinelli et al., 2024).

the treatment and control villages overlap, Table 2 reveals systematic differences in these distributions (they are unbalanced without adjustment) in the treatment and control villages.

However, the question still remains, whether conditioning on these pretreatment variables W_i's, i.e., the covariates, is sufficient for the independence of D_i with $Y_i(0)$ and $Y_i(1)$ in (1). This is our key assumption and, as is well-known, this cannot be validated statistically.

Therefore, we take the following approach in presenting our results on causal effects.

First, we will present in Section 4 our conclusions about the impact of *JS* by point-identifying and estimating the causal parameters of interest under a selection on observables assumption that is the conditional independence condition in (1). There does not seem to be an overlap-failure in our data (Table 2), and hence these estimated causal effects should not suffer from extrapolation-bias as long as (1) holds. Moreover, these estimated effects are stable across different choices of commonly used estimators, and they are also precise — both zero and non-zero effects are estimated precisely. This suggests that problems of limited overlap is not a pressing concern in our analysis (Khan and Tamer (2010)). Thus, our estimated causal effects are reliable provided that our key assumption (1) is true.

Then, in Section 5, we will revisit (1) and analyze the sensitivity of our conclusions (from Section 4) to departures from the conditional independence assumption in (1) using the methods of sensitivity analysis proposed by Masten and Poirier (2018, 2020) and Chernozhukov et al. (2024). Such departures signify the insufficiency of W_i in generating the required conditional independence in (1) such that the treatment is "as if random".

4 Empirical results: Causal effects

An overarching goal of *JS* is to promote open community discussion of the oppressive consequences of patriarchy and thereby reduce adherence to patriarchal norms and the beliefs they are embedded in. We test the hypothesis whether community-based participatory theater can shift the focus of attention in spousal violence from the manhood of the assailer

to his cruelty to his wife and make domestic violence socially unacceptable. We do this by estimating the impact of *JS* on various indicators of dominance – different types of spousal abuse, legitimacy of domestic violence, social participation in actions to prevent domestic violence and knowledge about legal recourse in case of incidents of domestic violence.

We discuss the regression adjustment (RA) estimates of the ATE ($E[Y_i(1) - Y_i(0)]$) and the baseline ($E[Y_i(0)]$) for all these outcomes. These results and also those for ATT and ATUT with their respective baseline estimates are reported in Tables 3 and 4. We also used alternative estimation methods– inverse probability weighting (IPW), combination of RA and IPW, and matching based on the observed covariates and propensity scores respectively, and double-debiased augmented IPW estimators based on covariate selections using lasso. The estimated effects were similar to those based on RA.[5] Supplemental Appendix A presents a subset of these results. As we noted earlier, one could always (more safely) interpret our results as the effects in our sample and not in any population (Imbens, 2004); our standard errors are then conservative, i.e., the effects are in fact more significant than reported.

4.1 Spousal Abuse

Acts of domestic violence (in the last 12 months) as reported by the wives: We estimate the impact of *JS* on physical abuse, sexual abuse, and wives who did not face any physical, sexual or emotional abuse — as reported by the female respondents. The definitions with detailed instructions to the field investigators on how to ask the questions are in Supplemental Appendix B. The National Family Health Survey 2015-16 (2017), the round that is the closest match to our data collection that started in 2014, reports that in rural West Bengal, 35.4 percent of married women in the age group 15-49 have experienced physical violence and 39 percent have experienced some form of domestic violence – physical,

[5]One might prefer IPW estimates over the RA estimates since in popular parlance one could label our discussion of confounding and identification as "design based" rather than "model based". Regardless, we discuss the RA estimates in the main text of the paper since regression is the standard econometric tool used in empirical development economics. However, we note that the IPW (and IPWRA) causal estimates and sensitivity analysis based on them, reported in Supplemental Appendix A, are very similar to the RA ones.

sexual or emotional. These statistics are roughly similar in our sample. We find the ATE of a village's exposure to *JS* is a reduction in the proportion of physically abused married women by 9.3 percentage points (p-value < 0.01) from the baseline ($E[Y_i(0)]$) of 32.6 percentage points (col. 1, Table 3). Exposure to *JS* reduced the proportion of sexually abused married women by close to half — 8.2 percentage points (p-value < 0.01) from the baseline of 17.8 percentage points (col. 2, Table 3). Finally, taking into account *any* form of spousal abuse – physical, sexual or emotional – a village's exposure to *JS* increased the proportion of abuse-free married women by 15.7 percentage points (p-value < 0.01) from the baseline of 45.8 percentage points (col. 3, Table 3).

***Legitimacy of Domestic Violence*:** Women are more likely to face domestic abuse due to acceptability of violence in the society as a norm and where an uneven relationship to domination and control by the husband over the woman exists. Field investigators asked women and their husbands whether a husband was justified to hit or beat his wife under any of the following circumstances: she goes out without telling him, neglects the children and household work, argues with him, refuses to have sex with him, disrespects her in-laws, is suspected of having an illicit relationship, and has not borne a son. Table 3 (col. 4) reports no meaningful effect of *JS* on this outcome using the wive's response. This is not surprising because the estimated baseline $E[Y_i(0)]$ of only 9 percentage points is already low. The effect of JS is only about a decrease of .7 percentage points, which although not small in relative terms (8 percent decrease from the baseline, and about 12 percent of the standard deviation of the observed Y_i), is not meaningfully large in absolute terms. By contrast, Table 4 (col. 1) shows that *JS* reduced the proportion of husbands who said that wife beating was justified in at least one of the aforementioned circumstances by 16.4 percentage points (p-value $< .01$) from the baseline of 21.9 percentage points. Interestingly, for this outcome, the estimated $E[Y_i(1)]$ for the husbands is lower than that for their wives (5.5 versus 8.3 percentage points), although the baseline $E[Y_i(0)]$ is not.

***Awareness that Domestic Violence Is Against the Law*:** India prohibited do-

mestic violence for the first time in 2006 through a comprehensive law – the Protection of Women from Domestic Violence (PDVA). The new law gives a female victim the right to file a police report and claim monetary compensation for medical treatment and loss of earnings, and also to claim maintenance from her husband for herself and her children. The law bars the husband and in-laws from evicting her from the marital home. However, laws designed to protect women from domestic violence are of no use if women are not aware of them. In rural India, awareness of such legislation reaching vulnerable women is less likely. Many *JS* plays therefore, as part of raising awareness among women about the PDVA, refer to women's rights under the 2006 law. The core *JS* team instructs the satellite teams to mention them whenever relevant in the interactive sessions with the audience. A village's exposure to *JS* reduced the proportions of wives and husbands who were not aware of the legal protections against domestic violence under the PDVA by 12.7 and 10.6 percentage points (p-value $<$ .01) from their respective baselines of 53.7 and 55.9 percentage points respectively (col. 5, Table 3 and col. 2, Table 4).

4.2 Voice in the household

***Empowerment of women*:** We also test whether *JS* empowered women in terms of participating in household decision making with their husbands. Field investigators asked each female respondent and, separately, her husband if she had not participated with him in making decisions in any of the 7 domains: education of the children, family health care, major household purchases, her visits to her relatives, the children's marriages, number of children to bear, and use of contraception. Based on the wife's response, a village's exposure to *JS* reduced this non-participation by almost half of the baseline; in particular, a reduction of 8.4 percentage points (p-value $<$ .01) from a baseline of 17.5 percentage points (col. 6, Table 3). By contrast, the husband's response to the same question about wife's non-participation in joint decision making is a reduction of 3 percentage points (p-value $<$ .01) from the baseline of 5.8 percentage points (col. 3, Table 4). The dissimilarity in the baseline for the same

non-participation variable based on the wife's and husband's responses suggests that the wife's non-participation, i.e., "voicelessness", is possibly differently perceived by the spouses — husbands are less likely to perceive it. Regardless, the *JS* intervention leads to significant reduction in "voicelessness" in both cases.

4.3 Engagement in community activities to alleviate violence against women

***Participation in Community Action to Report Acts of Domestic Violence witnessed as a bystander*:** Our survey asked the married women and their husbands if they would report to anyone acts of domestic violence that they witnessed as a bystander. A village's exposure to *JS* made almost universal the willingness of individuals to report incidents of domestic violence witnessed by them to someone, viz. villagers, members of husband's family, members of maternal family, police station, or gram panchayat (village council). *JS* increased reporting acts of domestic violence witnessed by women bystanders by 12.3 percentage points (p-value $< .01$) from an already high baseline of 84.3 percentage points (col. 7, Table 3). *JS* increased the proportion among men bystanders by 5.1 percentage points (p-value < 0.01) from an even higher baseline of 92.2 percentage points (col. 4, Table 4). The near universal willingness of married women and men, when treated, to report acts of domestic violence witnessed by them is consistent with the hypothesis that *JS* helped to change village norms by making wife beating a less normal behavior to villagers. The majority of female respondents preferred reporting an incident of domestic violence that they witnessed to formal institutions, such as the local police station or the village council. Less than 45 percent expressed their willingness to report the incident to a member of either their paternal or maternal family or to the villagers. Husbands of the female respondents did not show any preference in reporting to family members or villagers over institutions like the police station or the village council.

***Participate in community drive to destroy illegal liquor shops*:** Gender violence based on social norms and practices as explained above is further aggravated because of alcohol and substance abuse by the perpetrators (World Health Organization, 2014). For example, in the latest National Family Health Survey 2019-21 (2023) report on West Bengal, 85 percent of women experience spousal violence if their husband gets drunk often. The socio-economic burdens on families of excessive alcohol consumption are a central theme in some *JS* plays. We test whether *JS* encouraged villagers – both married men and women – to participate in community drives to destroy illegal liquor shops, whose presence is widely believed to contribute to domestic violence. Our survey asked, "If there was an illegal liquor shop in your community, would you participate in demolishing it?" We estimate that the proportion of married women in *JS* villages willing to participate in the demolition activities increased by 18.1 percentage points (p-value < 0.01) from the baseline of 46.2 percentage points (col. 7, Table 3). It also increased the husband's willingness to join community drives to demolish illegal liquor shops by 20.7 percentage points (p-value $< .01$) from the baseline of 70.1 percentage points (col. 5, Table 4). This makes it near universal willingness by men to participate in community drives to make their communities free of local illegal liquor shops.

It is however noteworthy that the baseline estimates for willingness in reporting domestic violence and in destruction of illegal liquor shops among men are much more than that among women, suggesting that there might be social desirability bias in the responses among men for these variables. This observation about social desirability in response is however less supported by the other responses of the men (cols. 1-3, Table 4).

4.4 Robustness check: Placebo/negative Outcomes

To appreciate the role of the selection on observables assumption in (1) based on which we address the selection bias due to *JS*' non-random treatment allocation, it is useful to consider outcomes on which there should be no causal effect of *JS*. Clean examples would be pretreatment values of the outcomes studied so far. We do not have the pretreatment

outcomes (else we would have included them in W_i in (1)). However, we know about households' posttreatment (2016) assets such as color TV. One might argue that JS should not affect ownership of color TV, i.e., this is a so-called placebo/negative outcome. (Results for ownership of refrigerator are qualitatively similar to that for color TV.)

If we run a naive regression of the ownership of color TV (proportion of households in a village that own color TV) on JS then we find a negative coefficient -.072 for JS with relatively high significance: p-value is 5.9%.

On the other hand, when we adjust for the covariates as suggested by our selection on observables assumption (1) and exactly in the same way as we did for all the outcomes discussed in our paper, we find that this significance vanishes both because the estimate gets much smaller and because the standard error gets bigger. For example, based on RA and IPW/IPWRA, we find the ATE estimates to be -.035 and -.045 respectively, and we find the associated the p-values to be 48.8% and 39.1% respectively; i.e., the significance vanished. (The results for ATT estimates based on RA and IPW/IPWA are even more insignificant with p-values 51.9% and 88.4% respectively.) Provided that one believes that JS should not affect ownership of color TV, it is reassuring to see that the adjustment for the covariates based on (1) used in our analysis seems to correct for the difference between the treatment and control villages and thereby avoid spurious inference of labeling a non-effect as an effect.

5 Sensitivity analysis of estimated causal effects

JS did not randomly select the villages but sought to perform in relatively populous villages that were geographically distant from towns, and not well connected to the outside world. Such villages were likely to have more conservative beliefs about gender roles. This caused selection problems in identifying our causal parameters of interest. For example, if JS targeted villages where physical abuse would be higher on average without an intervention, then this selection bias would be positive and thus the naïve comparison would underestimate

(biased toward zero) the average reduction in abuse due to the intervention on the ATT (e.g., p. 54, Angrist and Pischke (2009)). The sign of the selection bias for ATE is more challenging, and the results of naive regression do not provide bounds for ATE.

We addressed the selection problem by maintaining the selection on obervables condition (1); i.e., we assumed that conditional on a set of observed covariates W_i, the villages where *JS* performs were "as if" randomly chosen, in the sense that conditioning on W_i is sufficient to make the treatment D_i independent of the potential outcomes $Y_i(0)$ and $Y_i(1)$. Informed by our discussions with *JS*, for each village i, we postulated these observed confounders W_i to be the (log) population density in 1991 per kilometer square, (log) distance in kilometer to the nearest town in 1991, access to communication facilities in 1991, proportion of females among children under 6 years in 1991, and literacy growth rate between 1981 and 1991.

The selection on observables assumption in (1) cannot be validated statistically. However, we can analyze the sensitivity of our causal results to violations of (1). We do this following two distinct routes due to Masten and Poirier (2018)) and Chernozhukov et al. (2024).[6] We report the strength of unobserved confounders needed to overturn (change the sign of) our causal conclusions. It is ultimately up to the reader to decide if they believe that such strength of the unobserved confounder, as deduced by our sensitivity analysis, is likely.

Nevertheless, using standard benchmarking exercises, we find that an unobserved confounder needs to be nearly as strong as, or sometimes stronger than, *JS*'s stated observed selection criterion, a rather extreme case, to be able to overturn our conclusions. Such extreme requirements for the strength of unobserved confounders suggest that our causal results are insensitive to reasonable violations of our selection on observables condition.

[6]There are numerous important contributions to sensitivity analysis in economics; see, among many others, Imbens (2003), Altonji et al. (2005), Oster (2019), etc. We apply the sensitivity analysis due to Masten and Poirier (2018)) and Chernozhukov et al. (2024) for three reasons. First, as noted by their authors, these methods are valid under weak assumptions. Second, these methods work with bounded sensitivity parameters. Third, these methods were developed by economists or researchers with interest in economic applications and hence are probably more familiar to our readers. Various other novel methods of sensitivity analyses from the biostatistics, economics, epidemiology and statistics literature, when applied in our context, also led to the conclusion that the selection on observable assumption in (1) has to be incorrect in an unlikely big manner for our causal results to be overturned. These results are available from us.

5.1 Conditional c-dependence and breakdown points

Following Masten and Poirier (2018)), D_i is conditionally "c-dependent" on $Y_i(j)$ for $j = 0, 1$ given W_i provided that for all w and y in the support of W_i and $Y_i(j)$ respectively:

$$|P(D_i = 1|Y_i(j) = y, W_i = w) - P(D_i = 1|W_i = w)| \ \leq \ c \ \in [0, 1]. \tag{2}$$

$c = 0$ implies the conditional independence in our assumption (1). Larger c implies larger violation of (1), with the extreme case of $c \geq \max\{P(D_i = 0|W_i), P(D_i = 1|W_i)\}$ implying no restriction (zero-information) on $P(D_i = 1|Y_i(j) = y, W_i = w)$. Under the assumptions that D_i is conditionally c-dependent on $Y_i(j)$ for $j = 0, 1$ given W_i, Masten and Poirier (2018) provide the upper and lower bounds on ATE that are continuous and monotonic in c. For a given outcome with, e.g., ATE < 0, Masten and Poirier (2020) define the breakdown point c^* as the smallest c-dependence needed to overturn our ATE < 0, i.e.,

$$c^* = \inf_c \{\text{level of } c \text{ dependence such that the upper bound of ATE is} \geq 0\}$$

with the infimum defined as 1 if the set is empty. To assess if such a value c^* is reasonable, Masten and Poirier (2020)) recommend using the observed covariates to obtain the plausible c-dependence benchmarks as follows:

$$c_k = \sup_{w_k, w_{-k}} |P(D_i = 1|W_{i,k} = w_k, W_{i,-k} = w_{-k})) - P(D_i = 1|W_{i,-k} = w_{-k})|$$

where $W_{i,k}$ denotes the k-th covariate and $W_{i,-k}$ denotes the other covariates for different k's. c_k is the additional variation in the conditional probabilities of treatment due to the inclusion of $W_{i,k}$ in the conditioning set. Comparing with (2), we see that in the definition of c_k, the covariate $W_{i,k}$ takes the role of $Y_i(j)$, and the other covariates $W_{i,-k}$ take the role of W_i. These benchmarks c_k's are identifiable from the data, and when compared with the breakdown point c^*, they help us assess the plausibility of the violation of our selection on

observables assumption (1) that is big enough to overturn our causal conclusions.

Table 5 presents the breakdown point c^* for each outcomes studied in our paper using the "tesensitivity" package in STATA.[7] It also reports the elements $W_{i,k}$'s of the observed confounders in W_i, with c_k larger than c^*. Our causal conclusions will not survive, i.e., will change the sign, if the inclusion of $Y_i(j)$ in the conditioning set already consisting of W_i should have similar impact on the propensity score as the inclusion of $W_{i,k}$ in the conditioning set already consisting of $W_{i,-k}$. For most outcomes, including the key outcome of physical abuse, this requires $Y_i(j)$ to have an impact similar to the population density in the village i in 1991, or its distance from the district headquarter in 1991, that were in fact, by *JS*'s own account, the key variables in their selection of the treatment villages. That is, the unobserved confounder needs to be extremely strong in order to overturn our causal conclusions. We should also note that the breakdown point for the ATE on experiencing no abuse is 0.81. Hence, overturning the result that *JS* increased the proportion of households with no abuse requires $Y_i(j)$ (potential outcome from 2016) to have an impact on the propensity score of receiving treatment before 1998 that is much larger than that of all the observed determinants of the treatment, that we consider based on *JS*'s own account.

There are four outcomes for which our results are much more vulnerable to violations of the selection on observables condition in (1). The vulnerability is moderately high for the effects on the reduction in sexual abuse of the wife and the proportion of husbands who think that the wife does not participate in household decision making. On the other hand, the vulnerability is extremely high for the effects on the reduction in the proportion of wives who believe that beating the wife is justified, and the proportion of wives who think that the wife does not participate in household decision making. This fragility was expected a priori (even without any formal sensitivity analysis) given the small magnitude of the estimated effects. These are important outcomes by all accounts, and hence the failure of *JS* to have

[7] Masten et al. (2024) propose bootstrap inference for sensitivity analysis with c-dependence. As far as we know, this is so far neither included in their STATA package "tesensitivity" nor is a code publicly available separately. Since their bootstrap is nonstandard and is computationally intensive (see their Section 6.2), we do not do inference for analysis with c-dependence. We present inference for different analysis in Section 5.2.

robust impact on these four outcomes is certainly noteworthy.

The bounds for the ATE's at different values of c-dependence along with the benchmarks c_k's and the breakdown point c^* (intersection with the 0 effect line) are presented in Figures 2-14, using the STATA package "tesensitivity", for a visual display of this sensitivity analysis. This can be used by the reader to assess for themselves the plausibility of the strength of the unobserved confounders required to overturn our causal conclusions.[8]

5.2 Long and short models: Analysis with Reisz representor

Sensitivity analyses proposed by Altonji et al. (2005), Oster (2019), Cinelli and Hazlett (2020), etc. are well-suited for empirical applications that model homogeneous treatment effects, i.e., the so-called partially linear models. We allow for heterogeneous effects, i.e., D_i appears by itself and is also interacted with W_i in the regression adjustment. Hence, we instead follow Chernozhukov et al. (2024)'s approach to what they call the nonparametric model for ATE to obtain our second set of results on sensitivity analysis of (1).

Let U_i denote the possible unobserved confounder for the i-th village. Consequently, we will no longer maintain that the selection on observables assumption in (1) holds, i.e, the conditional independence in the so-called "short model" is true, and we will analyze how badly false it has to be in order for our causal conclusions to be overturned.

Nonparametric point-identification of ATE is ruled out if the short model is false. Under the assumption of conditional independence in the so-called "long model", i.e.:

$$Y_i(0) \text{ and } Y_i(1) \text{ are independent of } D_i \text{ conditional on } W_i \text{ and } U_i \text{ for all } i,$$

Chernozhukov et al. (2024) express the partially identified ATE as: $[\tau_L, \tau_U] = \tau \mp \sqrt{B^2}$

[8] Our outcome variables are proportions taking values in $[0, 1]$, and hence with large c the identified set for ATE should coincide with the no-information bound that should necessarily be of length 1 and should contain 0 effect. However, the 0 and/or 1 value of the outcomes are often far from the minimum or maximum of the observed outcomes in both the treatment and control villages (see Table 2). Hence, we do not use these implausibly conservative extremes but rather use the default setup of "tesensitivity" to obtain these bounds on the ATE's.

where $\tau = E\left[E[Y_i|D_i = 1, W_i] - E[Y_i|D_i = 0, W_i]\right]$ is the ATE under the short model (1),

$$B^2 = \left\{\eta^2_{(Y_i \sim U_i|D_i, W_i)}\right\} \times \left\{\frac{1 - \nu^2_{(D_i \sim U_i|W_i)}}{\nu^2_{(D_i \sim U_i|W_i)}}\right\} \times \left\{E\left[Var(Y_i|D_i, W_i)\right] E\left[\frac{1}{Var(D_i|W_i)}\right]\right\},$$

$$\eta^2_{(Y_i \sim U_i|D_i, W_i)} = 1 - \frac{E\left[Var(Y_i|D_i, W_i, U_i)\right]}{E\left[Var(Y_i|D_i, W_i)\right]} \text{ and } 1 - \nu^2_{(D_i \sim U_i|W_i)} = 1 - \frac{E\left[1/Var(D_i|W_i)\right]}{E\left[1/Var(D_i|W_i, U_i)\right]}.$$

$\eta^2_{(Y_i \sim U_i|D_i, W_i)}$ and $1 - \nu^2_{(D_i \sim U_i|W_i)}$ are the parameters for the sensitivity analysis of (1). (They are not identifiable from the data because of the presence of the unobserved confounder U_i.) $\eta^2_{(Y_i \sim U_i|D_i, W_i)} \in [0, 1]$ measures the additional gain in the explanatory power that U_i provides for Y_i, beyond what is already explained by D_i, W_i. $1 - \nu^2_{(D_i \sim U_i|W_i)} \in [0, 1]$ measures the relative gain in the average precision of the treatment D_i model due to U_i. $\eta^2_{(Y_i \sim U_i|D_i, W_i)} = 0$ and $1 - \nu^2_{(D_i \sim U_i|W_i)} = 0$ if "U_i is irrelevant" in the sense that the short model (1) is true. Then ATE is point-identified; else it is partially identified by $[\tau_L, \tau_U]$.

When e.g. $\widehat{\tau} < 0$, i.e., our estimated ATE < 0, then the sensitivity analysis studies the plausibility of the values of the sensitivity parameters $\left(\eta^2_{(Y_i \sim U_i|D_i, W_i)}, 1 - \nu^2_{(D_i \sim U_i|W_i)}\right)$ that make $\widehat{\tau}_U \geq 0$. In particular, in Table 6, we report the estimated Robustness Value which is the value $r_0 \in [0, 1]$ such that $\widehat{\tau}_U(r_0, r_0) = 0$ when $\widehat{\tau} < 0$ (alternatively, $\widehat{\tau}_L(r_0, r_0) = 0$ when $\widehat{\tau} > 0$). To account for the estimation error in $\widehat{\tau}$, we do a similar analysis with the upper 95% confidence bound of τ_U when $\widehat{\tau}_U < 0$; and the counterpart of the Robustness value in this context is denoted by Robustness Value (.05) and is defined as $r_0 \in [0, 1]$ such that the upper 95% confidence bound is 0 when $\eta^2_{(Y_i \sim U_i|D_i, W_i)} = 1 - \nu^2_{(D_i \sim U_i|W_i)} = r_0$.[9] When $\widehat{\tau}_L > 0$, we look instead at the lower 95% confidence bound of τ_L. The Robustness Value and the Robustness Value (.05) for each outcome are reported in Table 6. They suggest that our causal conclusions are able to survive (not change sign) even under rather large but

[9]While Chernozhukov et al. (2024) provide standard errors based on the semiparametrically efficient estimators, we will use nonparametric percentile bootstrap to obtain the confidence bounds. This is not controversial and we justify this as follows. We wish to report this sensitivity analysis with the RA, IPW and IPWRA estimators, since the results while similar are not identical. This means that the standard error based on the efficiency bound formula may be less appropriate with our small sample. On the other hand, formulae for standard errors based on parametric RA, IPW and IPWRA are tedious and different for these estimators. We use nonparametric bootstrap confidence bounds to bypass all these.

equal value of the sensitivity parameters. Results under "extreme scenarios" as in Cinelli and Hazlett (2020) (e.g., setting $\eta^2_{(Y_i \sim U_i | D_i, W_i)} = 1$ and studying sensitivity to $1 - \nu^2_{(D_i \sim U_i | W_i)}$), but with analysis like Chernozhukov et al. (2024), are available from us.

We provide a visual representation of this sensitivity analysis following Imbens (2003), Cinelli and Hazlett (2020) and Chernozhukov et al. (2024). Taking the sensitivity parameters $\left(\eta^2_{(Y_i \sim U_i | D_i, W_i)}, 1 - \nu^2_{(D_i \sim U_i | W_i)}\right)$ as the axes, we plot the contours of: (i) $\widehat{\tau}_U$ and the upper 95% confidence bound of τ_U if $\widehat{\tau} < 0$, and (ii) $\widehat{\tau}_L$ and the lower 95% confidence bound of τ_L if $\widehat{\tau} > 0$ for each outcome in Figures 15-26 with estimates based on regression adjustment. Contour plots with estimates based on IPW and IPWRA suggest similar sensitivity of ATE to violations of (1), and are presented in Figures 27-50 in Supplemental Appendix A.

How likely is it for the unobserved confounder U_i to have strength above the 0 contour line, and thus overturn our results? This is ultimately for each reader to decide based on their beliefs. The visual representation in Figures 15-26 may facilitate that decision.

Benchmarks to further facilitate assessing the plausibility of values of the sensitivity parameters that can overturn our causal conclusions from a non-inferential point of view (Figures 15-26 in panel (A)) and from an inferential point of view (Figures 15-26 in panel (B)) are obtained as follows. First, for each observed confounder $W_{i,k}$, and with $W_{i,-k}$ denoting the remaining elements of W_i as before, for $k = 1, \ldots, K$ $(K = 5)$ we estimate:

$$\eta^2_k \;=\; 1 - \frac{E\left[Var(Y_i | D_i, W_i)\right]}{E\left[Var(Y_i | D_i, W_{i,-k})\right]}, \quad \text{and} \quad 1 - \nu^2_k = 1 - \frac{E\left[1/Var(D_i | W_{i,-k})\right]}{E\left[1/Var(D_i | W_i)\right]}$$

resembling the original definitions but with $W_{i,k}$ playing the role of U_i and $W_{i,-k}$ playing the role of W_i. Then we take each of the K pairs $(\eta^2_k, 1 - \nu^2_k)$ for $k = 1, \ldots, K$ and also a more conservative worst-case measure $(\max_{1 \le k \le K} \eta^2_k, \max_{1 \le k \le K}(1 - \nu^2_k))$ as benchmarks for our sensitivity parameters. These six $(K + 1)$ points are plotted in the contour plots. The regular benchmarks are plotted in black while "the red dot" is the synthetic benchmark, i.e, the worst-case measure. These benchmarks are suggestive reference points for the readers about how much more influential the unobserved confounder U_i needs to be relative to the

observed confounders W_i in order to overturn our results on the ATE on each outcome.[10] For completeness of the representation, we also plot, marked in blue, the Robustness Value in the figures in panel (A) and the Robustness Value (.05) in the figures in panel (B).

As far as non-inferential results are concerned, i.e., if we are only looking at $\widehat{\tau}_U$ or $\widehat{\tau}_L$ (panel (a) in Figures 15-26), we observe that the unobserved confounder has to be quite influential to overturn many of the causal conclusion reported in the paper. We should emphasize that while weak covariates can artificially make the sensitivity analysis results look "good", our covariates include the population density and remoteness of the villages that were, by JS' own account, the key variables in their selection of the treatment villages. Therefore, based on what we learned from JS about their treatment assignment creteria, there does not seem to be any obvious candidate for unobserved confounder that could be strong enough to overturn these non-inferential results (panel (a) in Figures 15-26).

The inferential results (panel (b) in Figures 15-26), i.e., the upper 95% confidence bound for τ_U or the lower 95% confidence bound for τ_L, can be overturned by less extreme strength of the unobserved confounders. For example, the effects on the reduction in physical abuse and the reduction in the proportion of wives and husbands (respectively) that believe that the wife does not participate in joint decision making can be overturned if the unobserved component has strength as much as the worst-case synthetic benchmark. However, even then, an unobserved confounder that is as strong as the benchmarks, including the synthetic worst-case synthetic benchmark, cannot overturn the results for most of the outcomes. On the other hand, we now find that the effect on another outcome — the reduction in the proportion of husbands who believe that there is no law against domestic violence — to be extremely fragile.

The causal effects presented in our paper are noteworthy. However, since JS was not an RCT, these effects were obtained under the commonly used but strong assumption of

[10] We omit the outcome of response of married women to the question if physical abuse on them is justified. We saw in Section 4 that JS does not have any effect on this outcome even under (1), perhaps because the baseline proportion of women in our sample who justify such physical abuse is already low (estimated $E[Y_i(0)] = 9\%$). There is no meaningful effect on this outcome for the sensitivity analysis to make sense.

selection on observables in (1). Our extensive analysis in this section of the sensitivity of these estimated effects to departures from (1) reveal the following. We find that our estimates of the causal effects of JS are fragile for the four outcomes: the reduction in the proportion of wives who believe that beating the wife is justified, the reduction in the proportion of wives and husbands, respectively, who think that the wife does not participate in household decision making, and the reduction in the proportion of husbands who believe that there is no law against domestic violence. On the other hand, based on what we have learned from JS about their treatment assignment criteria, this sensitivity analysis does provide credibility to the effect of JS on the remaining outcomes.

6 Conclusion

From a policy perspective, we have established that participatory theater can be a very effective intervention. It may help individuals understand the world around them and how to change oppressive norms. Disadvantaged individuals may have internalized the norms of oppression, which makes it especially difficult for them to participate in changing them. Community-based Forum Theater has mechanisms to overcome these hurdles. It gives communities a platform for collectively finding new dramas that capture relationships that they have all experienced or observed. It gives disadvantaged people the power to imagine less oppressive relationships and to explore ways to bring them about. It thus gives disadvantaged groups the power to change cultural representations — an underestimated power that can bring about social change.

A performance begins by presenting from start to finish a short drama. The drama, which foregrounds women's suffering, causes some people to feel, often for the first time, outrage at the injustice. Then the drama starts again and a series of spect-actors, facilitated by the Joker, interrupt the drama and try to avert the oppression onstage. Their interactions may be a rehearsal for social change.

This paper is the first large-scale impact evaluation of Forum Theater, and *JS* is one of its largest practitioners. We estimate its impact on the prevalence and acceptability of domestic violence. *JS* did not randomly select the West Bengal villages where its troupes performed. We use the selection on observables assumption to identify and estimate the impact of village exposure to Forum Theater. This approach allows us to estimate the average treatment effect, average treatment effect on the treated villages, etc. of exposure to *JS* theater despite the fact that the intervention did not have a built-in evaluation module during its implementation. The effects are big. Their statistical significance is also strong in general and the use of conservative critical values to account for multiple testing by controlling for the family-wise error rate does not change the causal conclusions of our paper. We also report two sets of sensitivity analysis to study the robustness of our causal conclusions to potential violation of the selection on observables assumption. We find that big violations due to extremely strong unobserved confounder are required to overturn the effects that we observe in our analysis for most of the outcomes of interest.

We find that within at most 11 years of exposure to performances on themes of patriarchy and alcohol abuse, a community's tolerance of domestic violence and the proportion of couples in which domestic violence has occurred substantially decreased. The responses to our survey of both wives and their husbands support the conclusion that *JS* changed norms of a husband's proper behavior towards his wife.

Lawmakers, policy stakeholders, and civil society organizations in India have implemented several interventions to address domestic violence. These include legal reforms such as the Protection of Women from Domestic Violence Act (2005); support services like shelter homes, counselling, and crisis intervention centres (e.g., Sakhi Centres launched by the Ministry of Women and Child Development in 2015); and initiatives promoting women's education and economic empowerment, such as West Bengal's Kanyashree Prakalpa—a conditional cash transfer program launched in 2013 to delay child marriage, a known precursor to domestic violence. There are no Conditional Cash Transfer (CCT) schemes in India that

target reduction in domestic violence directly. CCT programs like Kanyashree aim to improve women's status by incentivizing education and delaying marriage, their effectiveness in reducing domestic violence remains uncertain. In rural West Bengal, data from both National Family Health Survey 2015-16 (2017) and National Family Health Survey 2019-21 (2023) show that the percentage of women aged 20–24 who were married before the age of 18 has remained unchanged at 42%. Furthermore, National Family Health Survey 2019-21 (2023) reports that 31.1% of ever-married women have experienced physical, sexual, or emotional abuse by their husbands. Moreover, the last figure likely understates the true scale of the problem, given pervasive underreporting due to patriarchal norms, social stigma, and limited trust in legal institutions. These statistics, call into question the scheme's long-term impact on child marriage and, by extension, domestic violence.

Globally, there are successful CCT programs like the Oportunidades/Prospera in Mexico (Bobonis and ands R. Castro, 2013), Give Directly in Kenya designed to alleviate extreme poverty (Haushofer and Shapiro, 2016) that reduce intimate partner violence. The notable differences in the program design between these programs and the CCTs in India are that their focus was on alleviating poverty together with changing norms and activism against gender inequality and violence against women rather than only focusing on keeping girls in school.

Mass media entertainment is another intervention that embeds educational messages in fictional stories (edutainment). It can influence some behaviours, e.g., consumption of iron-fortified salt (Banerjee et al., 2018), unsafe sex (Banerjee et al., 2019a), and reporting of corruption (Blair et al. (2019); and Cooper et al. (2020)). Yet edutainment has had little success in changing deep-seated preferences and attitudes towards domestic violence. Two evaluations have been conducted to date of the impact of edutainment on attitudes towards domestic violence (Banerjee et al. (2019b); and Cooper et al. (2020)). Both find limited impact on the belief that wife beating can be legitimate. To change an individual's core values about the morality of violence against women requires that he or she come to see

women as people with rights to make decisions about their lives. But if most people in a community accept a given prescriptive social norm, then even community members who do not share the beliefs on the grounds of those norms will have an incentive to follow it. The communication approach is clearly bottom-up, radically participatory, community based and led by the oppressed (Dutta (2011); Thomas and vad de Fliert (2015); and Tufte (2017)). *JanaSanskriti* in its interventions to change social norms does exactly that.

Economists increasingly recognize that the "scaffolds on which institutional structures rest" are beliefs that people hold to be true and reasonable and that they use to understand the social world (North (2005), pp. 48-49). An ideology can block the processing of information that would correct false beliefs about the inherent inequality of men and women. Our results show that Forum Theater undermined the scaffolds that support patriarchy. This outcome has eluded many other kinds of interventions—legal reforms, edutainment, information transmission, and mandated increases in female education. Through its direct effect on participants and communities, and their second-order effects on spectators' social networks, *JS* may end the self-perpetuating cycle of domestic violence in the villages in which it regularly performs.

Forum Theater is a package of many interventions that might be separated so that they can be more easily scaled up. Future research should implement experiments on ways to scale up elements of this package and address a wide range of social problems. What other social meanings could Forum Theater change? Possible targets are the social meanings that underpin the tolerance of corruption, corporal punishment of children, child marriage, discrimination, and resistance to policies to reduce global warming.

7 Bibliography

Agüero, J. M. (2019). Information and behavioral responses with more than one agent: The case of domestic violence awareness campaigns. (2019-04).

Akerlof, G. A. (2020). Sins of omission and the practice of economics. *Journal of Economic Literature*, 58(2):405–418.

Akerlof, R., N. M. and Rayo, L. (2020). Stories at work. *AEA Papers and Proceedings*, 110:199–214.

Altonji, J. G., E., T., and Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, 113: 151–184.

Ang, D. (2023). The birth of a nation: Media and racial hate. *American Economic Review*, 116:1424–1460.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Asante, E. and Zakaria, B. S. (2021). Participatory theatre in environmental health education: The woarabeba experience. *American Research Journal of Humanities and Social Sciences*, 12(4):121–128.

Banerjee, A., Barnhardt, S., and Duflo, E. (2018). Can iron-fortified salt control anemia? evidence from two experiments in rural bihar. *Journal of Development Economics*, 133(C):127–146.

Banerjee, A., La-Ferrara, E., and Orozco, V. (2019a). The entertaining way to behavioral change: Fighting hiv with mtv. NBER Working Paper No. 26096.

Banerjee, A., La-Ferrara, E., and Orozco, V. (2019b). Entertainment, education, and attitudes toward domestic violence. *American Economic Review*, 109:133–137.

Benabou, R., Falk, A., and Tirole, J. (2018). Narratives, imperatives and moral persuasion. NBER Working Paper No. 24798.

Blair, G., Littman, R., and Paluck, E. L. (2019). Motivating the adoption of new community-minded behaviors: An empirical test in nigeria. *Science Advances*, 5(3):eaau5175.

Boal, A. (1979). *Theatre of the Oppressed*. Urizen, New York.

Bobonis, G. J. and ands R. Castro, M. G.-B. (2013). Public Transfers and Domestic Violence: The Roles of Private Information and Spousal Control. *American Economic Journal: Economic Policy*, 5:179–205.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge University Press, New York.

Census of India (1991). Primary census abstracts.

Chernozhukov, C., Cinelli, C., Newey, W., Sharma, A., and Syrgkanis, V. (2024). Long Story Short: Omitted Variable Bias in Causal Machine Learning. Working Paper.

Chowdhry, P. (2015). Popular perceptions of masculinity in rural north indian oral traditions. *Asian Ethnography*, 74(1):5–36.

Cinelli, C., Forney, A., and Pearl, J. (2024). A Crash Course in Good and Bad Controls. *Sociological Methods and Research*, 53: 1071–1104.

Cinelli, C. and Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1):39–67.

Cooper, J., Green, D., and Wilke, A. (2020). Reducing violence against women in uganda through video dramas: A survey experiment to illuminate causal mechanisms. *American Economic Review*, 110:615–619.

Derne, S. (1994). Hindu men talk about controlling women: Cultural ideas as a tool of the powerful. *Sociological Perspectives*, 37(2):203–227.

Doepke, M., Tertilt, M., and Voenna, A. (2012). The economics and politics of women's rights. *Annual Review of Economics*, 4:339–372.

Dutta, M. J. (2011). *Communicating Social Change: Structure, Culture, and Agency*. Routledge.

Egeland, B. (1993). *A history of abuse is a major risk factor for abusing in the next generation*, pages 197–208. Sage Publications, Newbury Park, CA.

Esposito, E., Rotesi, T., Saia, A., and Thoenig, M. (2023). Reconciliation Narratives: The Birth of a Nation after the US Civil War. *Ameican Economic Rebiew*, 113:1461–1504.

Freire, P. (1970). *Pedagogy of the Oppressed*. Herder and Herder, New York.

Ganguly, S. (2010). *Jana Sanskriti: Forum Theatre and Democracy in India*. Routledge, Oxon.

Ganguly, S. (2017). *From Boal to Jana Sanskriti*. Routledge, London.

Haushofer, J. and Shapiro, J. (2016). The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya Get access Arrow. *The Quarterly Journal of Economics*, 131:1973–2042.

Hoff, K. and Stiglitz, J. (2010). Equilibrium fictions: A cognitive approach to societal rigidity. *American Economic Review*, 100:141–146.

Htun, M. and Jensenius, F. R. (2020). Violence against women: Laws, norms, and challenges ahead. *Daedalus*, 149(1):144–159.

Imbens, G. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review*, 93: 126–132.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Review of Economics and Statistics*, 86:4–29.

Jha, S. and Shayo, M. (2019). Valuing peace: The effects of financial market exposure on votes and political attitudes. *Econometrica*, 87:1561–1588.

Kahan, D. M., Peters, E., Dawson, E. C., and Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioral Public Policy*, 1(1):54–86.

Khan, S. and Tamer, E. (2010). Irregular Identification, Support Conditions, and Inverse Weight Estimation. *Econometrica*, 78:2021–2042.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Masten, M. A. and Poirier, A. (2018). Identification of treatment effects under conditional partial independence. *Econometrica*, 86(1):317–351.

Masten, M. A. and Poirier, A. (2020). Inference on breakdown frontiers. *Quantitative Economics*, 11(1):41–111.

Masten, M. A., Poirier, A., and Zhang, L. (2024). Assessing Sensitivity to Unconfoundedness: Estimation and Inference. *Journal of Business and Economic Statistics*, 42: 1–13.

Mills, S. (2009). Theatre for transformation and empowerment: a case study of jana sanskriti theatre of the oppressed. *Development in Practice*, 19(4-5):550–559.

Mohan, D. (2010). *Scripting Power: Jana Sanskriti, On and Offstage*. Camp, Kolkata.

Narayan, D., Patel, R., Schafft, K., Rademacher, A., and Koch-Schulte, S. (2000). *Voices of the poor. Can anyone hear us?* Oxford University Press, Oxford.

National Family Health Survey 2015-16 (2017). National family health survey 2015-16.

National Family Health Survey 2019-21 (2023). National family health survey 2019-21.

North, D. C. (2005). *Understanding the Process of Economic Change*. Princeton University Press, Princeton, NJ.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37: 187–204.

Scharinger, J. (2013). Participatory theater, is it really? a critical examination of practices in timor-leste. *ASEAS - Austrian Journal of South-East Asian Studies*, 6(1):102–119.

Shiller, R. (2019). *Narrative Economics*. Princeton University Press, Princeton, NJ.

Thambu, N., Shanmugavelu, G., and Mahayudin, Z. (2020). The use of forum theatre as creative pedagogy in teaching and learning of moral education in secondary schools in malaysia. *EPRA International Journal of Multidisciplinary Research*, 6(8):2455–3662.

Thomas, P. and vad de Fliert, E. (2015). *Interrogating the Theory and Practice of Communication for Social Change : The Basis for a Renewal*. Palgrave Macmillan The Limited.

Tufte, T. (2017). *Communication and Social Change: A Citizen Perspective*. Polity.

World Health Organization (2014). *Global Status Report on Alcohol and Health, 2014*. WHO Press, Geneva.

Yarrow, R. (2012). Performing agency. *Studies in South Asian Film and Media*, 4(2):211–226.

Yarrow, R. (2017). From performers to spectactivists: Jana sanskriti's training for agency in and beyond theatre. *Indian Theatre Journal*, 1(1):29–37.

8 Appendix: Tables and Figures

Table 1: Cross-country statistics on women justifying domestic violence and experiencing intimate partner violence, 2017

		Percentage of women who consider a husband is justified in hitting/beating his wife for at least one specified reason	Percentage of women who ever suffered violence from an intimate partner			Percentage of women who consider a husband is justified in hitting/beating his wife for at least one specified reason	Percentage of women who ever suffered violence from an intimate partner
	COUNTRY	(1)	(2)		**COUNTRY**	(3)	(4)
(A) LOW-INCOME COUNTRIES	Guinea	92.1	80	**(D) HIGH-INCOME COUNTRIES**	Kuwait	37	NA
	Afghanistan	80.2	60.8		Chinese Taipei	21.6	NA
	South Sudan	78.5	NA		Germany	19.6	22
	Somalia	75.7	NA		Korea	18.4	16.5
	Congo (Dem Rep)	74.8	50.7		Switzerland	15.2	9.8
	Chad	73.5	28.6		Argentina	11.6	NA
	Burundi	72.9	46.7		United States	11	35.6
	Mali	72.6	34.6		Chile	10.3	6.7
	Ethiopia	63	28		United Kingdom	10.2	29
	Niger	59.6	NA		Sweden	10.2	28
	Tajikistan	59.6	20.3		Spain	9.6	13
	Haiti	58.9	20.8		Japan	8.9	15.4
	Uganda	58.3	49.9		Hungary	8.7	21
	Tanzania	58	41.7		Poland	7.9	13
	Senegal	56.5	78		Canada	7.8	1.9
	Yemen	48.7	67		France	6.6	26
	Madagascar	45.2	30		Netherlands	6.4	25
	Burkina Faso	43.5	11.5		Italy	5.3	19
	Nepal	42.9	25		Australia	3.2	16.9
	Rwanda	41.4	34.4		Austria	3	13
	Zimbabwe	38.7	35.4		Belgium	2	24
	Benin	36	68.6		Czech Republic	2	21
	Togo	28.7	22.1		Greece	2	19
	Mozambique	22.9	21.7		Portugal	2	19
	Malawi	16.3	37.5				
	MEDIAN	58.3	35		**MEDIAN**	8.8	19
(B) LOWER MIDDLE-INCOME COUNTRIES	Sri Lanka	53.2	16.6	**(C) UPPER MIDDLE-INCOME COUNTRIES**	South Africa	61.2	20.6
	Myanmar	51.2	33		Iraq	54.8	21.2
	Cambodia	50.4	20.9		Guinea	52.6	56.9
	Côte d'Ivoire	47.9	25.9		Algeria	48.2	NA
	Zambia	46.9	42.7		Malaysia	41.5	NA
	Pakistan	42.2	85		China	32.7	NA
	Kenya	41.8	39.4		Peru	32.2	33.2
	Uzbekistan	41.5	NA		Azerbaijan	28	13.5
	Cameroon	36.1	51.1		Ecuador	25.2	37.5
	Egypt	35.7	NA		Russia	23.3	19.6
	Nigeria	34.7	16.2		Iran	21	66
	Indonesia	34.5	18.3		Jordan	18	23.6
	Sudan	34	NA		Kazakhstan	14.2	16.5
	Bangladesh	28.3	53.3		Turkey	13.3	38
	Ghana	28.3	24.4		Colombia	11.1	37.4
	Viet Nam	28.2	34.4		Guatemala	11	18
	Angola	25.2	34.8		Thailand	8.6	44.2
	India	22.1	28.7		Brazil	8.5	33.5
	Morocco	22	30		Romania	7.5	24
	Tunisia	18.6	20.3		Mexico	5	14.1
	Bolivia	16.1	64.1		Belarus	4.1	25
	Philippines	12.9	16.9		Cuba	3.9	NA
	Honduras	12.4	21.6		Serbia	3.8	23.7
	Ukraine	2.9	13.2		Dominican Republic	2	20.4
	MEDIAN	34.25	28.7		**MEDIAN**	16.1	23.85

Notes. The respondents in are nationally representative samples of women 15-49 years of age. The five reasons that are specified are: the wife burns the food, she argues with her husband, she goes out without telling him, she neglects the children, and she refuses sexual relations with him. Cols. 2 and 4 show the percentage of women who have ever suffered intimate partner physical and/or sexual violence. Omitted from the table are countries with populations less than 0.1% of the world's total population and countries for which data are not available. *Source.* OECD Gender, Institutions, and Development Database (GID-DB), 2019. https://stats.oecd.org/Index.aspx?DataSetCode=GIDDB2019

Table 2: Descriptive statistics

	Control Villages (61)					*Treatment Villages (31)*				
Variables	**Mean**	**Median**	**Std Dev**	**Min**	**Max**	**Mean**	**Median**	**Std Dev**	**Min**	**Max**
Selection on observable covariates										
Population density in 1991 (in logs)	2.425	2.406	0.544	1.208	4.023	1.782	1.740	0.443	0.706	3.117
Distance to nearest town in 1991 (in logs)	2.829	2.833	0.717	1.386	3.912	3.782	4.094	0.520	2.485	4.605
Literacy growth rate between 1981 & 1991	0.067	0.061	0.081	-0.214	0.296	0.085	0.083	0.091	-0.152	0.285
Fraction of girls among 5 and under in 1991	0.498	0.490	0.032	0.406	0.593	0.491	0.486	0.031	0.429	0.574
Access to communication facilities in 1991	0.230	0.000	0.424	0	1	0.097	0.000	0.301	0	1
Wife's responses: Outcomes										
Whether physically abused in last 12 months	0.314	0.294	0.124	0.062	0.586	0.236	0.237	0.079	0.080	0.458
Whether sexually abused in last 12 months	0.185	0.176	0.087	0.000	0.360	0.135	0.120	0.071	0.031	0.306
Whether emotionally abused in last 12 months	0.354	0.353	0.130	0.129	0.692	0.286	0.280	0.081	0.133	0.453
Did not face any abuse	0.480	0.485	0.121	0.161	0.742	0.580	0.594	0.099	0.367	0.781
Physical abuse of wives is justified	0.093	0.086	0.065	0.000	0.296	0.058	0.057	0.042	0.000	0.156
No knowledge about laws on domestic violence	0.534	0.542	0.116	0.308	0.840	0.380	0.379	0.062	0.288	0.533
No joint decisions	0.181	0.185	0.084	0.000	0.417	0.110	0.089	0.091	0.000	0.500
Would report acts of domestic violence	0.854	0.871	0.091	0.581	1.000	0.956	0.960	0.032	0.879	1.000
Would demolish illegal liquor shops	0.469	0.462	0.139	0.083	0.800	0.679	0.688	0.115	0.233	0.837
Husband's responses: Outcomes										
Physical abuse of wives is justified	0.230	0.226	0.120	0.034	0.481	0.075	0.052	0.074	0.000	0.268
No knowledge about laws on domestic violence	0.555	0.559	0.113	0.250	0.913	0.418	0.407	0.149	0.172	0.867
No joint decisions	0.062	0.048	0.055	0.000	0.200	0.021	0.016	0.027	0.000	0.111
Would report acts of domestic violence	0.929	0.938	0.061	0.760	1.000	0.974	0.978	0.034	0.867	1.000
Would demolish illegal liquor shops	0.723	0.720	0.140	0.250	0.944	0.908	0.920	0.065	0.667	1.000

Source: Census of India, 1991 and Primary survey, 2014-15.

Table 3: **Wife's responses (based on RA)**

	Impact of *JS*' interventions on individual actions						**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse	Sexual abuse	Abuse-free relationship	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATE E[Y(1) – Y(0)]	-0.0930 (0.025)	-0.0815 (0.020)	0.1565 (0.032)	-0.0072 (0.012)	-0.1266 (0.025)	-0.0843 (0.026)	0.1016 (0.018)	0.1809 (0.045)
ATT E[Y(1) – Y(0)|D=1]	-0.1128 (0.034)	-0.0273 (0.022)	0.1637 (0.033)	-0.0278 (0.016)	-0.1615 (0.030)	-0.0540 (0.025)	0.1515 (0.025)	0.2327 (0.042)
ATUT E[Y(1) – Y(0)|D=0]	-0.0830 (0.028)	-0.1091 (0.026)	0.1529 (0.040)	0.0033 (0.014)	-0.1089 (0.029)	-0.0997 (0.032)	0.0762 (0.018)	0.1545 (0.057)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.0780 (0.021)	-0.0504 (0.017)	0.1002 (0.024)	-0.0346 (0.011)	-0.1546 (0.019)	-0.0713 (0.019)	0.1022 (0.013)	0.2101 (0.027)
Expected potential outcome without treatment: E[Y(0)]	0.3255 (0.019)	0.1774 (0.012)	0.4581 (0.018)	0.0906 (0.009)	0.5365 (0.017)	0.1754 (0.012)	0.8375 (0.015)	0.4615 (0.021)
Expected potential outcome with treatment: E[Y(1)]	0.2325 (0.017)	0.0959 (0.016)	0.6147 (0.027)	0.0834 (0.009)	0.4099 (0.018)	0.0911 (0.023)	0.9392 (0.009)	0.6423 (0.040)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.3138 (0.016)	0.1852 (0.011)	0.4796 (0.016)	0.0929 (0.008)	0.5342 (0.015)	0.1813 (0.011)	0.8542 (0.012)	0.4691 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.3486 (0.031)	0.1621 (0.018)	0.4160 (0.028)	0.0861 (0.014)	0.5411 (0.028)	0.1639 (0.019)	0.8048 (0.024)	0.4465 (0.036)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.2308 (0.023)	0.0761 (0.023)	0.6325 (0.038)	0.0962 (0.012)	0.4254 (0.025)	0.0815 (0.030)	0.9304 (0.013)	0.6236 (0.054)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.2358 (0.014)	0.1349 (0.013)	0.5797 (0.018)	0.0583 (0.008)	0.3796 (0.011)	0.1100 (0.016)	0.9564 (0.006)	0.6792 (0.021)

Standard errors in parentheses. No cell color indicates significance at $\leq$ 1%, indicates significance at 1% $< x \leq$ 5% and indicates significance at $>$ 5%.

Table 4: **Husband's responses (based on RA)**

	Impact of *JS*' interventions on individual actions			**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)
ATE E[Y(1) – Y(0)]	-0.1644 (0.022)	-0.1060 (0.033)	-0.0305 (0.010)	0.0510 (0.011)	0.2074 (0.028)
ATT E[Y(1) – Y(0)|D=1]	-0.1232 (0.030)	-0.1487 (0.037)	-0.0296 (0.011)	0.0655 (0.014)	0.2245 (0.032)
ATUT E[Y(1) – Y(0)|D=0]	-0.1854 (0.025)	-0.0843 (0.048)	-0.0310 (0.011)	0.0436 (0.011)	0.1986 (0.031)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.1549 (0.020)	-0.1365 (0.030)	-0.0407 (0.009)	0.0448 (0.010)	0.1849 (0.021)
Expected potential outcome without treatment: E[Y(0)]	0.2192 (0.017)	0.5590 (0.015)	0.0578 (0.007)	0.9219 (0.008)	0.7099 (0.019)
Expected potential outcome with treatment: E[Y(1)]	0.0548 (0.014)	0.4531 (0.030)	0.0273 (0.007)	0.9729 (0.007)	0.9173 (0.020)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.2299 (0.015)	0.5549 (0.015)	0.0616 (0.007)	0.9289 (0.008)	0.7233 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.1981 (0.027)	0.5671 (0.025)	0.0504 (0.010)	0.9081 (0.013)	0.6836 (0.030)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.0445 (0.020)	0.4707 (0.045)	0.0306 (0.009)	0.9725 (0.009)	0.9219 (0.026)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.0750 (0.013)	0.4184 (0.027)	0.0209 (0.005)	0.9736 (0.006)	0.9081 (0.012)

Standard errors in parentheses. No cell color indicates significance at $\leq$ 1%, indicates significance at $1\% < x \leq 5\%$ and indicates significance at $> 5\%$.

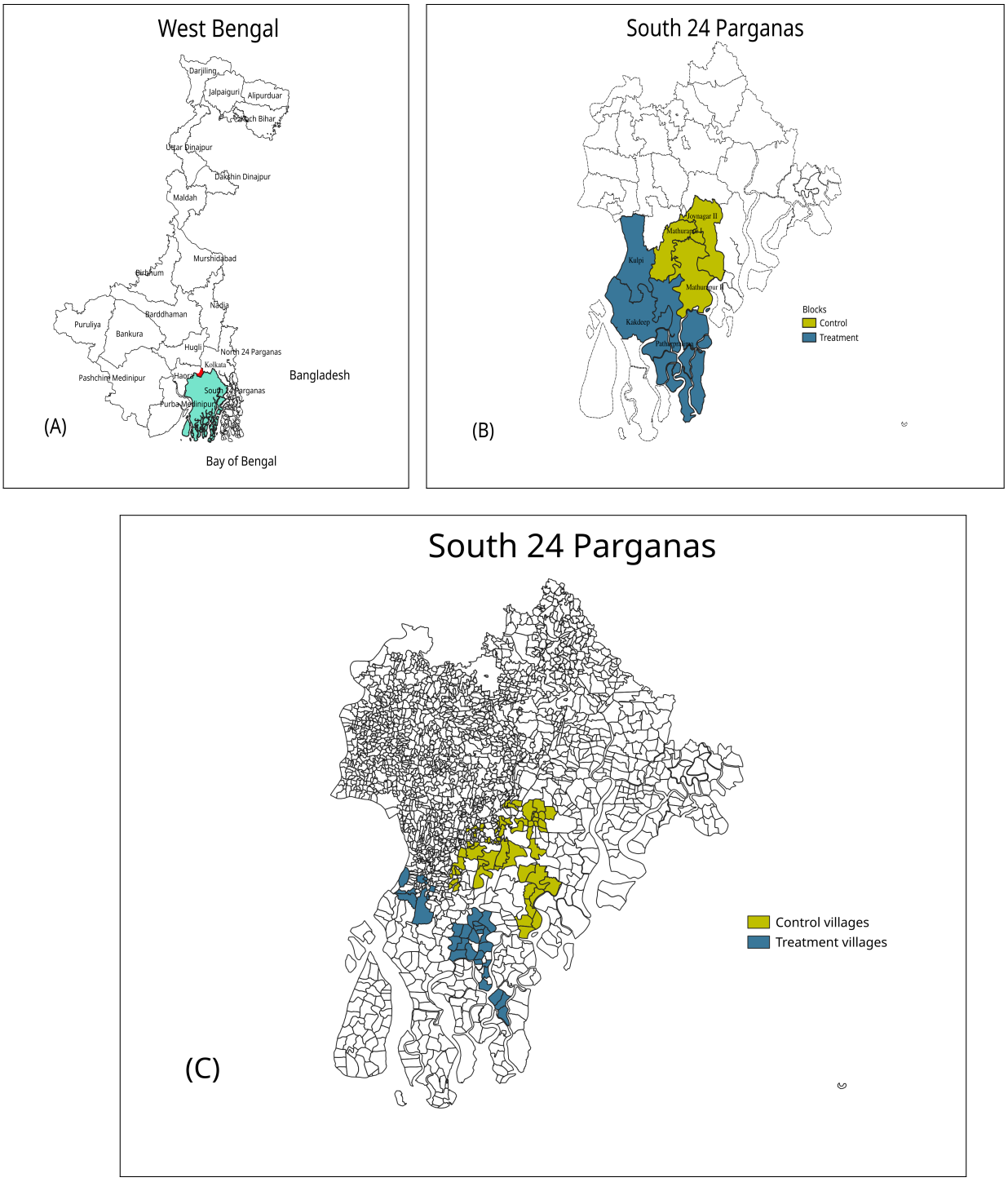
Table 5: Masten-Poirier Breakdown points

Outcomes	Breakdown point	Observed Confounders (1991)
Wife's responses		
Physical abuse	0.4234	Distance to nearest town
Sexual abuse	0.0641	Literacy growth rate (1981-91), Population density, Distance to nearest town
Abuse-free relationship	0.8077	
Physical abuse by husband is justified	0.0288	Literacy growth rate (1981-91), Population density, Distance to nearest town
No knowledge about PDVA	0.2628	Population density, Distance to nearest town
No joint decisions	0.0304	Literacy growth rate (1981-91), Population density, Distance to nearest town
Report acts of domestic violence	1.0000	
Demolish illegal liquor shops	0.2532	Population density, Distance to nearest town
Husband's responses		
Physical abuse by husband is justified	0.2885	Population density, Distance to nearest town
No knowledge about PDVA	0.3013	Population density, Distance to nearest town
No joint decisions	0.0705	Literacy growth rate (1981-91), Population density, Distance to nearest town
Report acts of domestic violence	0.1987	Population density, Distance to nearest town
Demolish illegal liquor shops	0.3654	Population density, Distance to nearest town

Table 6: Robustness values of the short estimates

	RA		IPW		IPWRA	
Outcomes	Robustness Value	Robustness Value (0.05)	Robustness Value	Robustness Value (0.05)	Robustness Value	Robustness Value (0.05)
Wife's responses						
Physical abuse	0.29	0.14	0.25	0.1	0.29	0.14
Sexual abuse	0.31	0.12	0.26	0.16	0.27	0.07
Abuse-free relationship	0.42	0.22	0.36	0.22	0.39	0.16
Physical abuse by husband is justified	0.04	0.01	0.1	0.01	0.06	0.01
No knowledge about PDVA	0.37	0.2	0.41	0.28	0.38	0.23
No joint decisions	0.31	0.1	0.29	0.16	0.29	0.1
Report acts of domestic violence	0.41	0.28	0.46	0.34	0.45	0.32
Demolish illegal liquor shops	0.4	0.17	0.42	0.28	0.39	0.15
Husband's responses						
Physical abuse by husband is justified	0.44	0.29	0.41	0.3	0.41	0.26
No knowledge about PDVA	0.28	0.11	0.36	0.24	0.32	0.16
No joint decisions	0.21	0.09	0.25	0.14	0.22	0.1
Report acts of domestic violence	0.31	0.18	0.29	0.16	0.3	0.14
Demolish illegal liquor shops	0.47	0.3	0.43	0.32	0.43	0.26

Figure 1: **Administrative areas of the survey.** The highlighted areas show (A) the district of South 24 Parganas, (B) control and treatment blocks of South 24 Parganas, and (C) control and treatment villages



Masten-Poirier bounds (wife's responses)

Figure 2: Physical abuse of wife

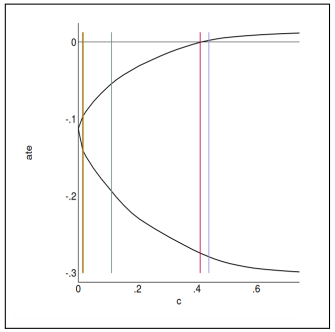


Figure 3: Sexual abuse of wife

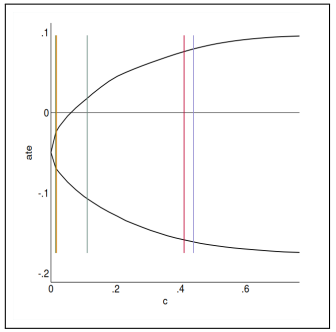


Figure 4: Abuse-free relationship

Figure 5: Wife beating by husband is justified

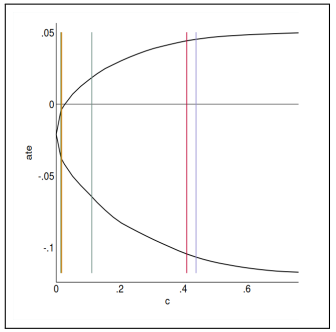


Figure 6: Wife does not know about the PDVA

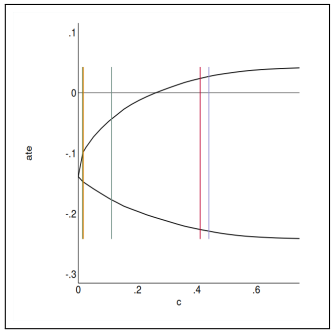


Figure 7: Wife has no voice in household

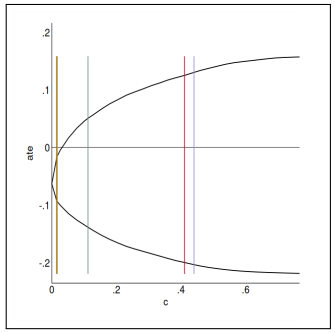


Figure 8: Report acts of domestic violence witnessed by her

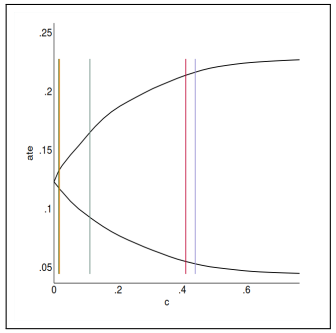
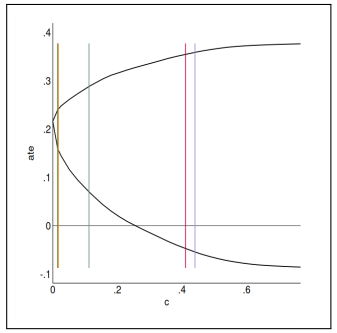
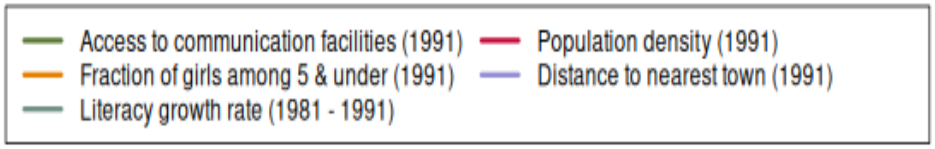


Figure 9: Demolish illegal liquor shops in the community





Masten-Poirier bounds (husband's responses)

Figure 10: **Wife beating by husband is justified**

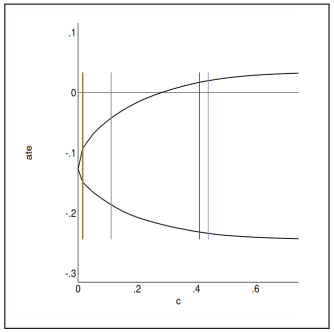


Figure 11: **Husband does not know about the PDVA**

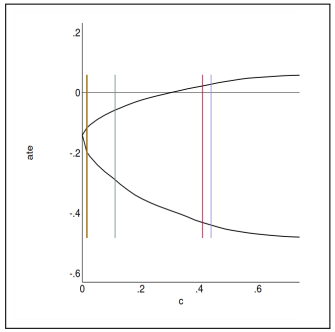


Figure 12: **Wife has no voice in household**

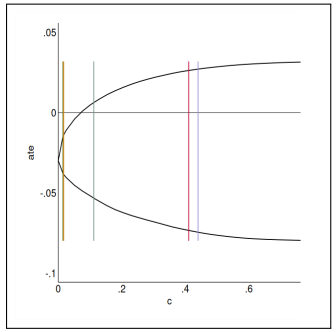


Figure 13: **Report domestic violence acts witnessed by him**

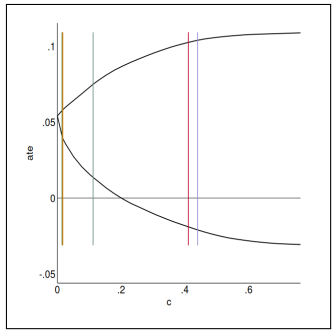
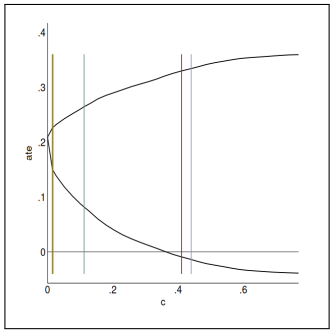


Figure 14: **Demolish illegal liquor shops in the community**



Contour plots based on Chernozhukov et al. (2024) based on RA (wife's responses)

Figure 15: Physical abuse of wife

A. Upper bound

B. Upper 95 % confidence bound for upper bound

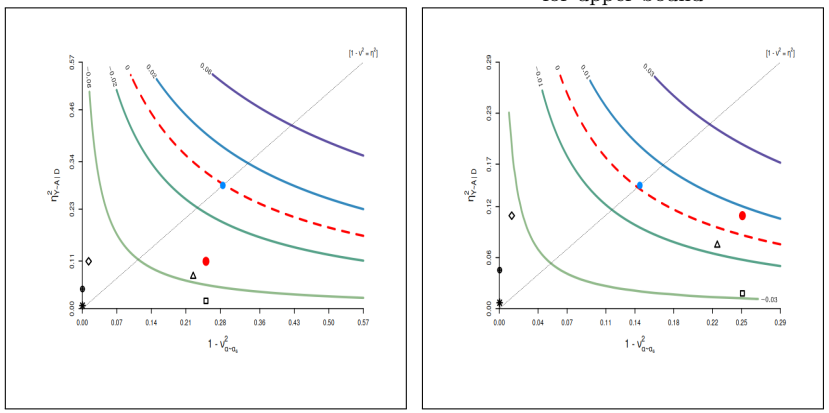


Figure 16: Sexual abuse of wife

A. Upper bound

B. Upper 95 % confidence bound for upper bound

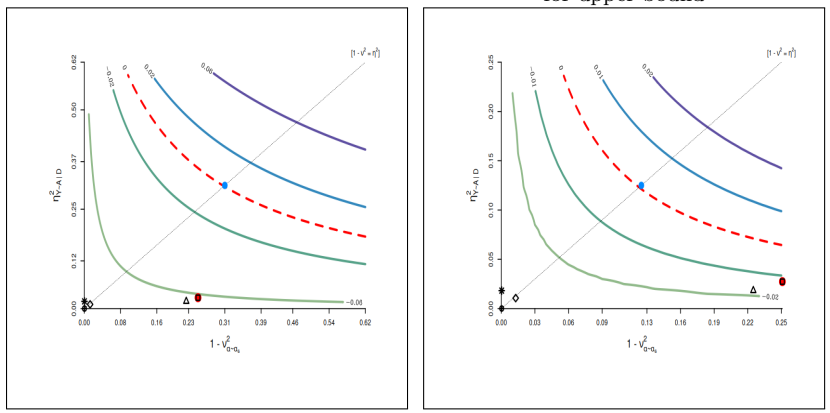
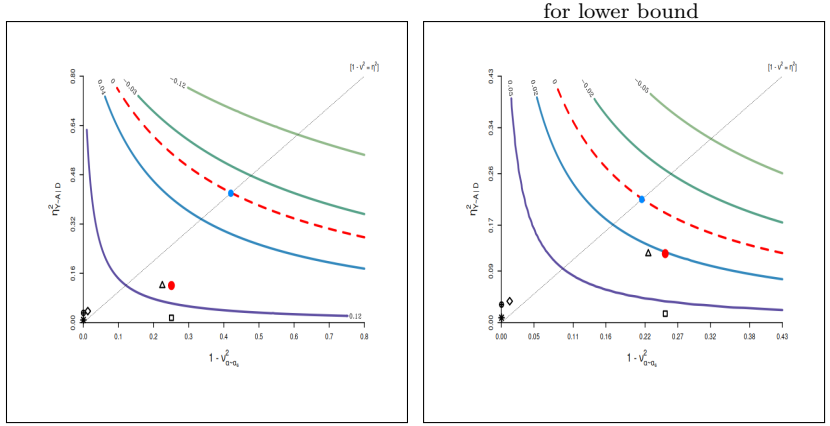


Figure 17: Abuse-free relationship

A. Lower bound

B. Lower 95 % confidence bound for lower bound



Contour plots based on Chernozhukov et al. (2024) based on RA (wife's responses)

Figure 18: **Wife does not know about the PDVA**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

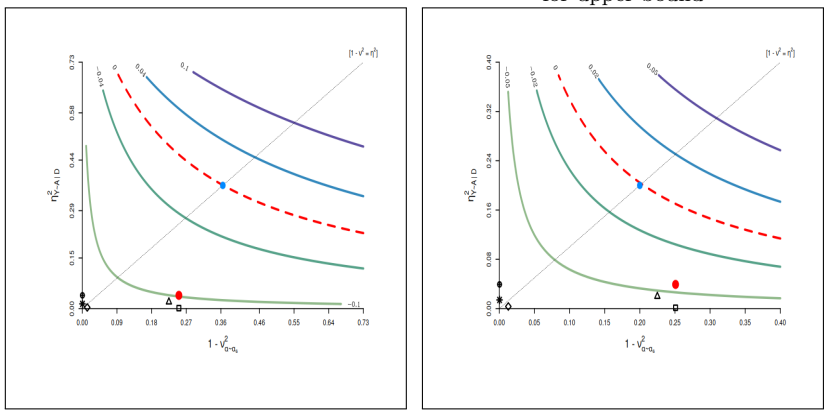


Figure 19: **Wife has no voice in household**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

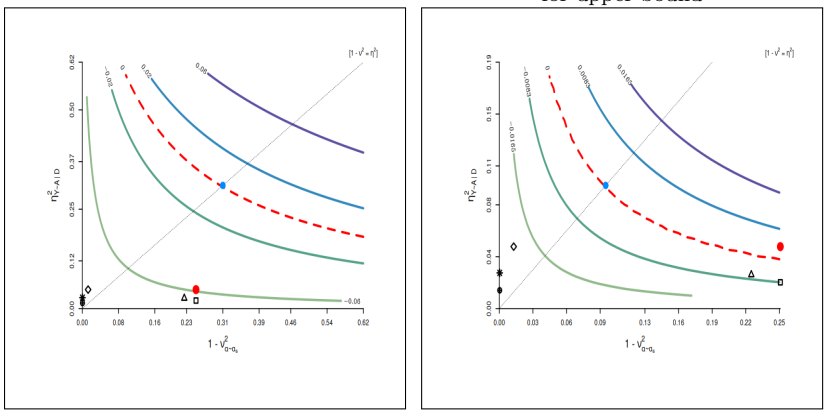


Figure 20: **Report domestic violence acts witnessed by her**

A. Lower bound

B. Lower 95 % confidence bound for lower bound

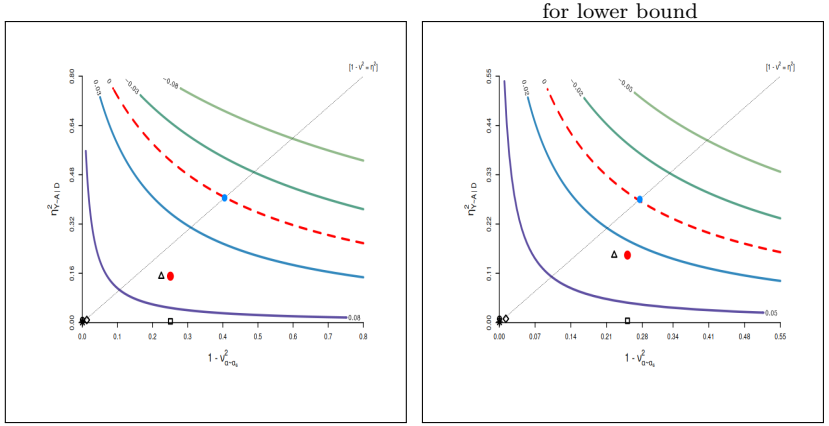
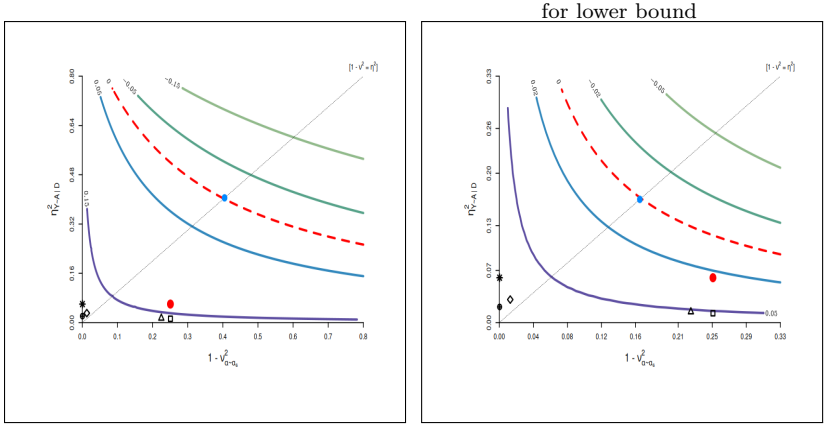
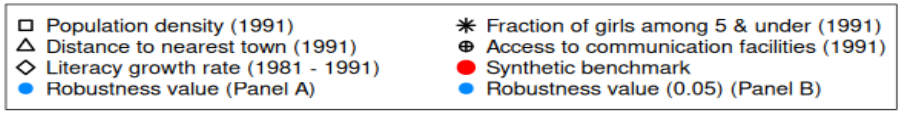


Figure 21: **Demolish illegal liquor shops in the community**

A. Lower bound

B. Lower 95 % confidence bound for lower bound





Contour plots based on Chernozhukov et al. (2024) based on RA (husband's responses)

Figure 22: **Wife beating by husband is justified**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

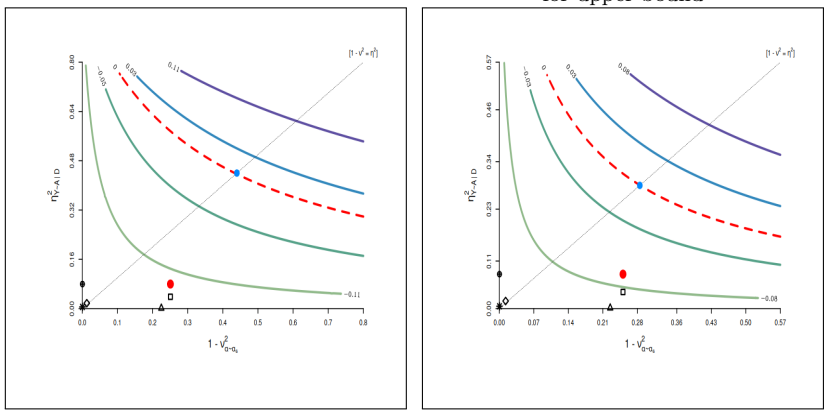


Figure 23: **Husband does not know about the PDVA**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

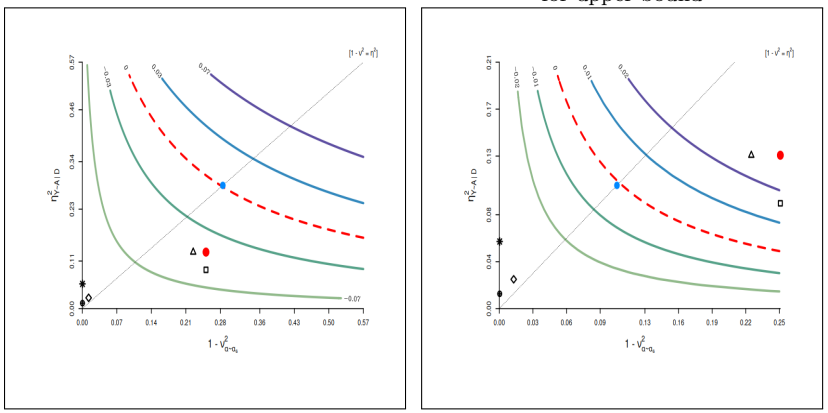


Figure 24: **Wife has no voice in household**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

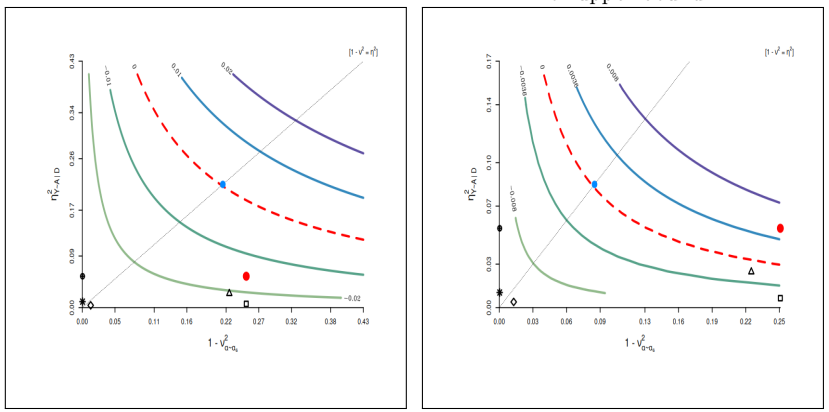
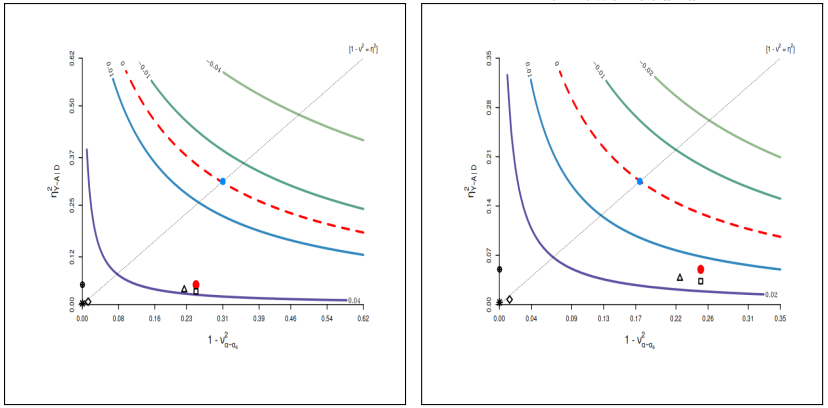


Figure 25: **Report domestic violence acts witnessed by him**

A. Lower bound

B. Lower 95 % confidence bound for lower bound



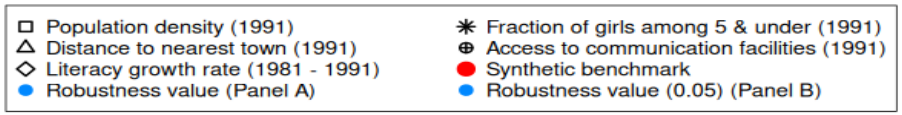
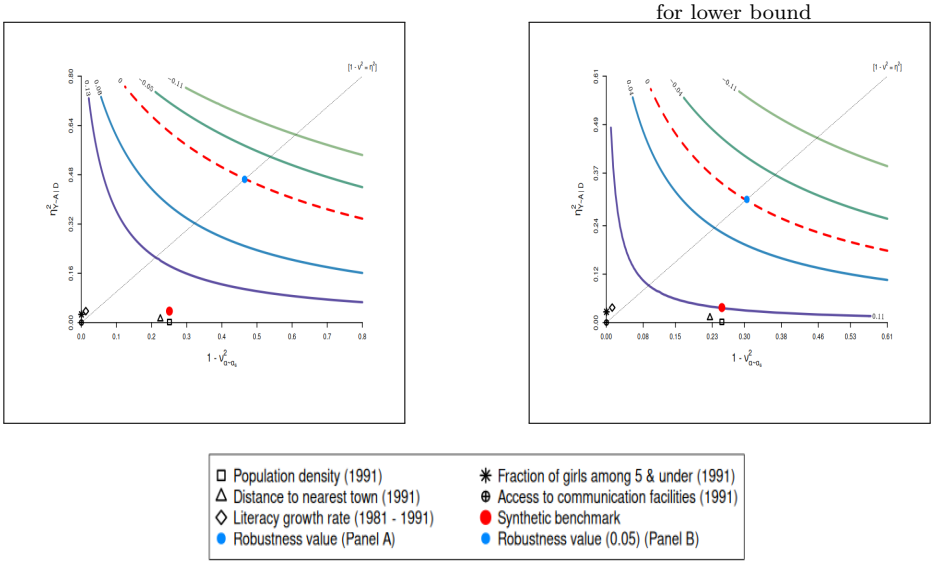


Figure 26: **Demolish illegal liquor shops in the community**

A. Lower bound

B. Lower 95 % confidence bound for lower bound



Supplemental Appendix:
Participatory Theater Empowers Women: Evidence from West Bengal, India

S. Chaudhuri[11] K. Hoff[12] J. Jalan[13] S. Santra[14]

Table of Contents

[11]Department of Economics, McGill University & Cireq, Montreal. Email: saraswata.chaudhuri@mcgill.ca.
[12]Department of Economics, Columbia University. Email: kh2830@columbia.edu.
[13]Centre for Studies in Social Sciences, and Centre for Training and Research in Public Finance and Policy, Kolkata. Email: jjalan@cssscal.org.
[14]Centre for Training and Research in Public Finance and Policy, Kolkata. Email: sattwik@cssscal.org.

A Results using alternative estimators

In the main text we reported the causal estimates based on regression adjustment (RA) use the "teffects" command in STATA with the "ra" option because regression is the default option in empirical economics. Although perhaps causal estimates based on inverse probability weighting (IPW) or IPWRA would have been more obvious choice since the motivation of our confounders was "design based" and not "model based". Therefore, exactly in the same pattern as Tables 3-4 from the main text, we will report here the causal results based on IPW (Tables 7-8) and IPWRA (Tables 9-10). Matching and augmented IPW estimates are similar and not reported for brevity.

Then, exactly in the same pattern as Figures 15-26 from the main text that were based on RA, here we will present the contour plots for the Chernozhukov et al. (2024)-type sensitivity analysis along with the benchmarks in Figures 27-38 and 39-50 for sensitivity analysis based on IPW and IPWRA respectively. Various other types of sensitivity analysis results are also available from us.

Table 7: **Wife's responses (based on IPW)**

	Impact of *JS*' interventions on individual actions						**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse	Sexual abuse	Abuse-free relationship	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATE E[Y(1) – Y(0)]	-0.0797 (0.030)	-0.0656 (0.016)	0.1272 (0.026)	-0.0172 (0.010)	-0.1491 (0.021)	-0.0780 (0.016)	0.1197 (0.018)	0.1941 (0.029)
ATT E[Y(1) – Y(0)|D=1]	-0.1062 (0.029)	-0.0230 (0.019)	0.1689 (0.027)	-0.0270 (0.011)	-0.1554 (0.021)	-0.0447 (0.023)	0.1721 (0.037)	0.1896 (0.029)
ATUT E[Y(1) – Y(0)|D=0]	-0.0685 (0.039)	-0.0856 (0.017)	0.1125 (0.030)	-0.0109 (0.011)	-0.1458 (0.027)	-0.0930 (0.015)	0.0993 (0.014)	0.1942 (0.034)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.0780 (0.021)	-0.0504 (0.017)	0.1002 (0.024)	-0.0346 (0.011)	-0.1546 (0.019)	-0.0713 (0.019)	0.1022 (0.013)	0.2101 (0.027)
Expected potential outcome without treatment: E[Y(0)]	0.3216 (0.016)	0.1776 (0.010)	0.4605 (0.016)	0.0908 (0.008)	0.5345 (0.014)	0.1739 (0.010)	0.8348 (0.017)	0.4748 (0.017)
Expected potential outcome with treatment: E[Y(1)]	0.2420 (0.025)	0.1121 (0.011)	0.5877 (0.020)	0.0736 (0.006)	0.3853 (0.015)	0.0959 (0.011)	0.9545 (0.005)	0.6689 (0.023)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.3138 (0.016)	0.1852 (0.011)	0.4796 (0.016)	0.0929 (0.008)	0.5342 (0.015)	0.1813 (0.011)	0.8542 (0.012)	0.4691 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.3421 (0.026)	0.1578 (0.013)	0.4109 (0.021)	0.0853 (0.009)	0.5350 (0.017)	0.1547 (0.014)	0.7843 (0.037)	0.4896 (0.024)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.2453 (0.036)	0.0996 (0.013)	0.5921 (0.026)	0.0820 (0.007)	0.3884 (0.023)	0.0882 (0.010)	0.9534 (0.006)	0.6633 (0.031)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.2358 (0.014)	0.1349 (0.013)	0.5797 (0.018)	0.0583 (0.008)	0.3796 (0.011)	0.1100 (0.016)	0.9564 (0.006)	0.6792 (0.021)

Standard errors in parentheses. No cell color indicates significance at $\leq 1\%$, indicates significance at $1\% < x \leq 5\%$ and indicates significance at $> 5\%$.

Table 8: Husband's responses (based on IPW)

	Impact of *JS*' interventions on individual actions			**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)
ATE E[Y(1) – Y(0)]	-0.1522 (0.020)	-0.1421 (0.022)	-0.0363 (0.008)	0.0468 (0.011)	0.1865 (0.019)
ATT E[Y(1) – Y(0)|D=1]	-0.0776 (0.034)	-0.1297 (0.032)	-0.0359 (0.011)	0.0614 (0.017)	0.1675 (0.025)
ATUT E[Y(1) – Y(0)|D=0]	-0.1838 (0.016)	-0.1480 (0.027)	-0.0360 (0.010)	0.0407 (0.012)	0.1948 (0.021)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.1549 (0.020)	-0.1365 (0.030)	-0.0407 (0.009)	0.0448 (0.010)	0.1849 (0.021)
Expected potential outcome without treatment: E[Y(0)]	0.2085 (0.017)	0.5530 (0.013)	0.0602 (0.006)	0.9243 (0.008)	0.7281 (0.017)
Expected potential outcome with treatment: E[Y(1)]	0.0563 (0.008)	0.4110 (0.017)	0.0239 (0.006)	0.9710 (0.007)	0.9146 (0.010)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.2299 (0.015)	0.5549 (0.015)	0.0616 (0.007)	0.9289 (0.008)	0.7233 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.1526 (0.030)	0.5481 (0.017)	0.0568 (0.011)	0.9122 (0.016)	0.7406 (0.024)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.0460 (0.007)	0.4069 (0.022)	0.0256 (0.008)	0.9696 (0.010)	0.9181 (0.012)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.0750 (0.013)	0.4184 (0.027)	0.0209 (0.005)	0.9736 (0.006)	0.9081 (0.012)

Standard errors in parentheses. No cell color indicates significance at $\leq$ 1%, indicates significance at $1\% < x \leq 5\%$ and indicates significance at $> 5\%$.

Table 9: **Wife's responses (based on IPWRA)**

	Impact of *JS*' interventions on individual actions						**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse	Sexual abuse	Abuse-free relationship	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATE E[Y(1) – Y(0)]	-0.0936 (0.021)	-0.0686 (0.014)	0.1423 (0.024)	-0.0096 (0.009)	-0.1339 (0.018)	-0.0805 (0.016)	0.1207 (0.019)	0.1758 (0.030)
ATT E[Y(1) – Y(0)|D=1]	-0.0979 (0.036)	-0.0121 (0.022)	0.1894 (0.039)	-0.0189 (0.010)	-0.1456 (0.020)	-0.0285 (0.023)	0.2108 (0.039)	0.1798 (0.030)
ATUT E[Y(1) – Y(0)|D=0]	-0.0829 (0.021)	-0.0890 (0.014)	0.1124 (0.022)	-0.0055 (0.010)	-0.1299 (0.017)	-0.1051 (0.013)	0.0940 (0.012)	0.1729 (0.030)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.0780 (0.021)	-0.0504 (0.017)	0.1002 (0.024)	-0.0346 (0.011)	-0.1546 (0.019)	-0.0713 (0.019)	0.1022 (0.013)	0.2101 (0.027)
Expected potential outcome without treatment: E[Y(0)]	0.3240 (0.018)	0.1748 (0.011)	0.4529 (0.017)	0.0894 (0.008)	0.5353 (0.015)	0.1707 (0.011)	0.8267 (0.018)	0.4726 (0.018)
Expected potential outcome with treatment: E[Y(1)]	0.2304 (0.012)	0.1062 (0.009)	0.5952 (0.016)	0.0798 (0.006)	0.4014 (0.011)	0.0902 (0.011)	0.9474 (0.006)	0.6484 (0.025)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.3138 (0.016)	0.1852 (0.011)	0.4796 (0.016)	0.0929 (0.008)	0.5342 (0.015)	0.1813 (0.011)	0.8542 (0.012)	0.4691 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.3338 (0.033)	0.1470 (0.017)	0.3903 (0.034)	0.0772 (0.010)	0.5252 (0.017)	0.1385 (0.015)	0.7456 (0.037)	0.4994 (0.025)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.2309 (0.015)	0.0962 (0.007)	0.5920 (0.015)	0.0874 (0.006)	0.4044 (0.009)	0.0762 (0.006)	0.9482 (0.004)	0.6420 (0.023)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.2358 (0.014)	0.1349 (0.013)	0.5797 (0.018)	0.0583 (0.008)	0.3796 (0.011)	0.1100 (0.016)	0.9564 (0.006)	0.6792 (0.021)

Standard errors in parentheses. No cell color indicates significance at $\leq 1\%$, indicates significance at $1\% < x \leq 5\%$ and indicates significance at $> 5\%$.

Table 10: **Husband's responses (based on IPWRA)**

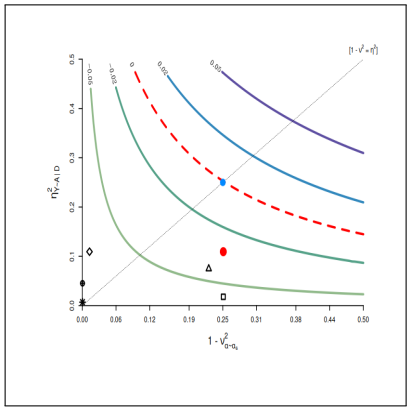
	Impact of *JS*' interventions on individual actions			**Impact of *JS*' interventions on individual's social consciousness**	
	Physical abuse by husband is justified	No knowledge about PDVA	No joint decisions	Report acts of domestic violence	Demolish illegal liquor shops
ESTIMAND	(1)	(2)	(3)	(4)	(5)
ATE E[Y(1) – Y(0)]	-0.1531 (0.019)	-0.1230 (0.021)	-0.0326 (0.007)	0.0504 (0.010)	0.1863 (0.019)
ATT E[Y(1) – Y(0)|D=1]	-0.0354 (0.038)	-0.1167 (0.031)	-0.0278 (0.011)	0.0647 (0.023)	0.1450 (0.025)
ATUT E[Y(1) – Y(0)|D=0]	-0.1919 (0.017)	-0.1327 (0.018)	-0.0356 (0.008)	0.0449 (0.008)	0.1892 (0.021)
Naïve regression E[Y(1)|D=1] - E[Y(0)|D=0]	-0.1549 (0.020)	-0.1365 (0.030)	-0.0407 (0.009)	0.0448 (0.010)	0.1849 (0.021)
Expected potential outcome without treatment: E[Y(0)]	0.2040 (0.017)	0.5534 (0.013)	0.0573 (0.006)	0.9228 (0.009)	0.7255 (0.017)
Expected potential outcome with treatment: E[Y(1)]	0.0509 (0.007)	0.4304 (0.016)	0.0247 (0.004)	0.9731 (0.004)	0.9118 (0.010)
Expected potential outcome without treatment for untreated: E[Y(0)|D=0]	0.2299 (0.015)	0.5549 (0.015)	0.0616 (0.007)	0.9289 (0.008)	0.7233 (0.018)
Expected potential outcome without treatment for treated: E[Y(0)|D=1]	0.1103 (0.034)	0.5351 (0.014)	0.0487 (0.010)	0.9089 (0.021)	0.7632 (0.024)
Expected potential outcome with treatment for untreated: E[Y(1)|D=0]	0.0380 (0.006)	0.4223 (0.010)	0.0259 (0.005)	0.9738 (0.004)	0.9125 (0.011)
Expected potential outcome with treatment for treated: E[Y(1)|D=1]	0.0750 (0.013)	0.4184 (0.027)	0.0209 (0.005)	0.9736 (0.006)	0.9081 (0.012)

Standard errors in parentheses. No cell color indicates significance at $\leq$ 1%, indicates significance at $1\% < x \leq 5\%$ and indicates significance at $> 5\%$.

Contour plots based on Chernozhukov et al. (2024) based on IPW (wife's responses)

Figure 27: Physical abuse of wife

A. Upper bound



B. Upper 95 % confidence bound for upper bound

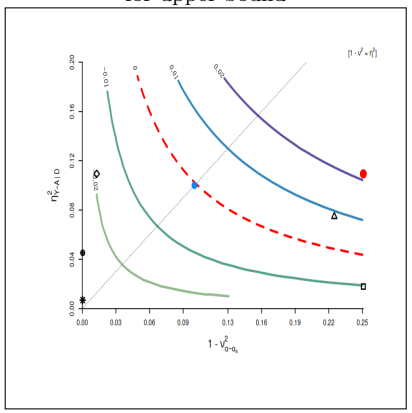
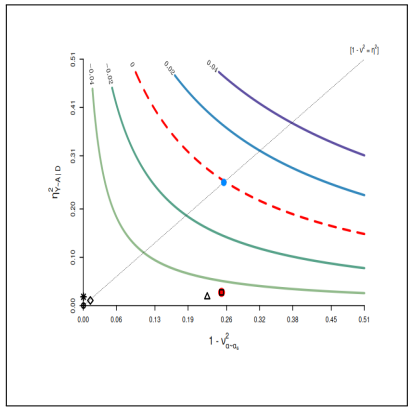


Figure 28: Sexual abuse of wife

A. Upper bound



B. Upper 95 % confidence bound for upper bound

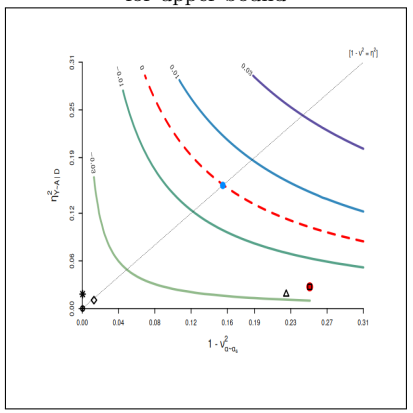
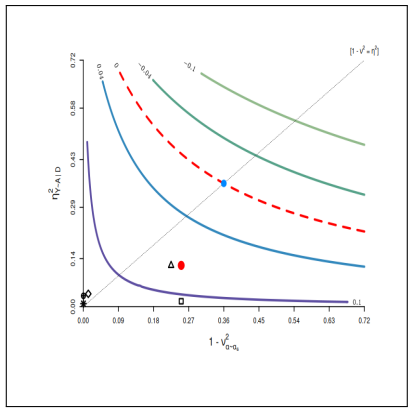
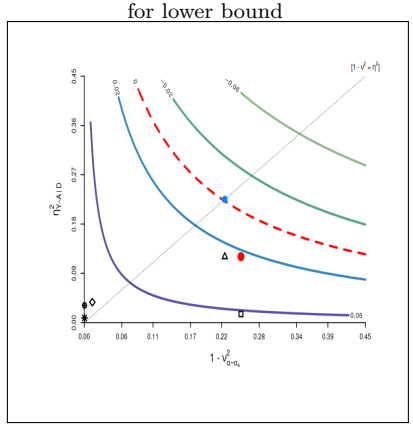


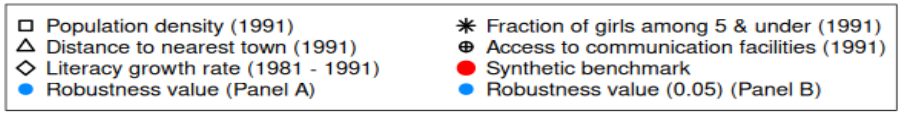
Figure 29: Abuse-free relationship

A. Lower bound



B. Lower 95 % confidence bound for lower bound

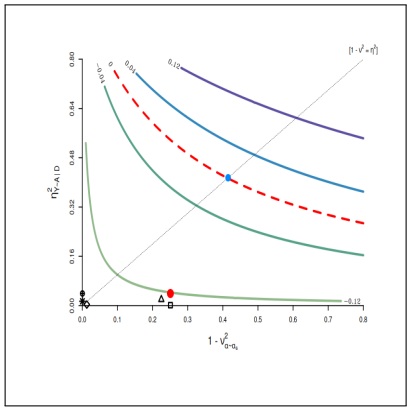




Contour plots based on Chernozhukov et al. (2024) based on IPW (wife's responses)

Figure 30: **Wife does not know about the PDVA**

A. Upper bound



B. Upper 95 % confidence bound for upper bound

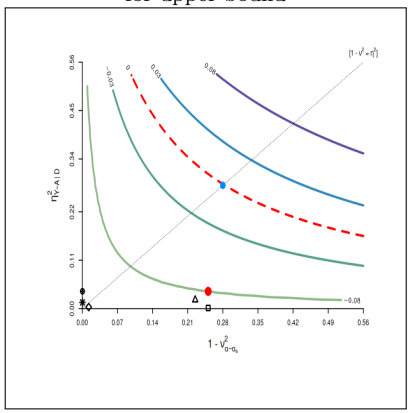
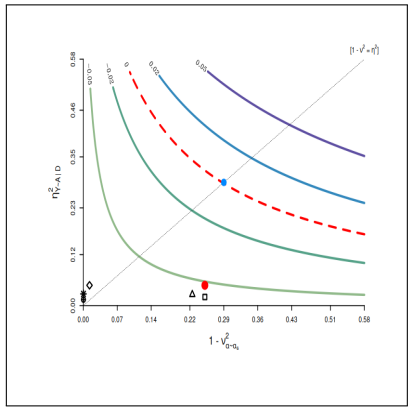


Figure 31: **Wife has no voice in household**

A. Upper bound



B. Upper 95 % confidence bound for upper bound

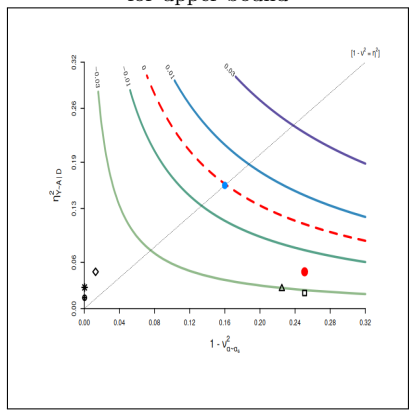
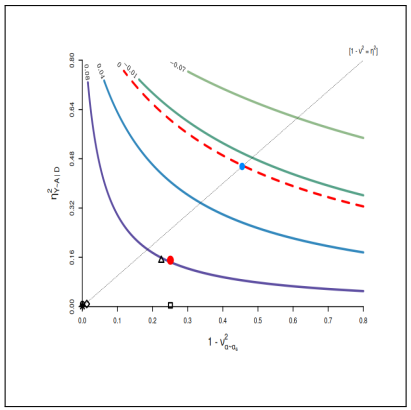


Figure 32: **Report domestic violence acts witnessed by her**

A. Lower bound



B. Lower 95 % confidence bound for lower bound

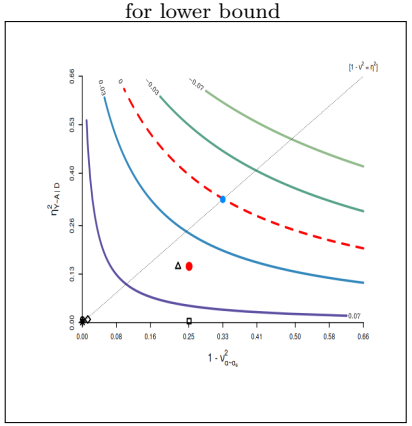
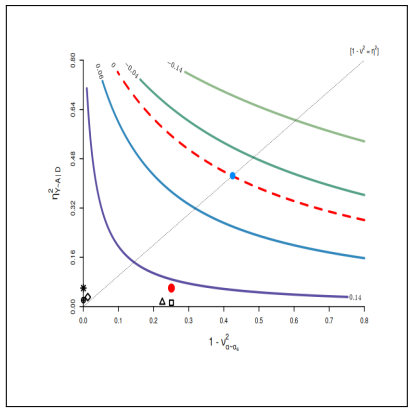
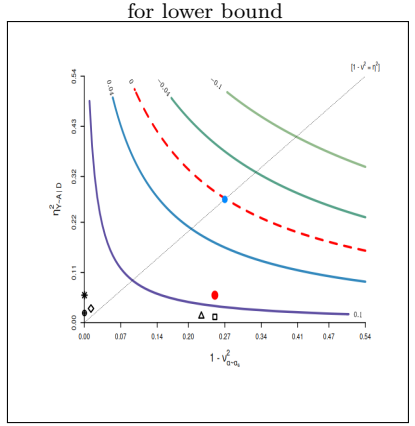


Figure 33: **Demolish illegal liquor shops in the community**

A. Lower bound



B. Lower 95 % confidence bound for lower bound



- Population density (1991)
- Distance to nearest town (1991)
- Literacy growth rate (1981 - 1991)
- Robustness value (Panel A)
- Fraction of girls among 5 & under (1991)
- Access to communication facilities (1991)
- Synthetic benchmark
- Robustness value (0.05) (Panel B)

Contour plots based on Chernozhukov et al. (2024) based on IPW (husband's responses)

Figure 34: **Wife beating by husband is justified**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

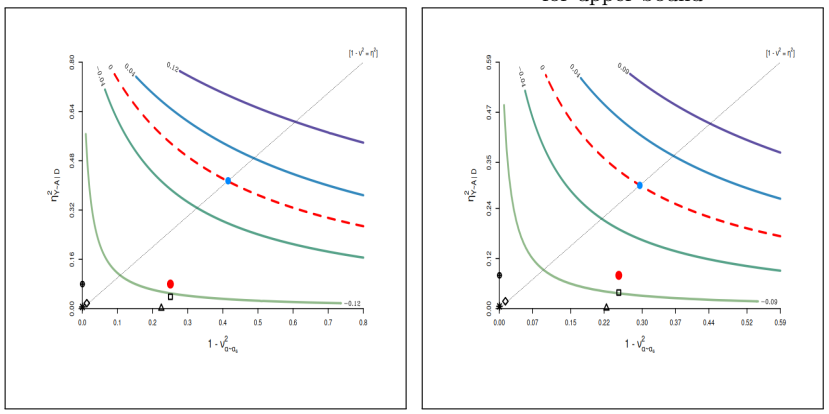


Figure 35: **Husband does not know about the PDVA**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

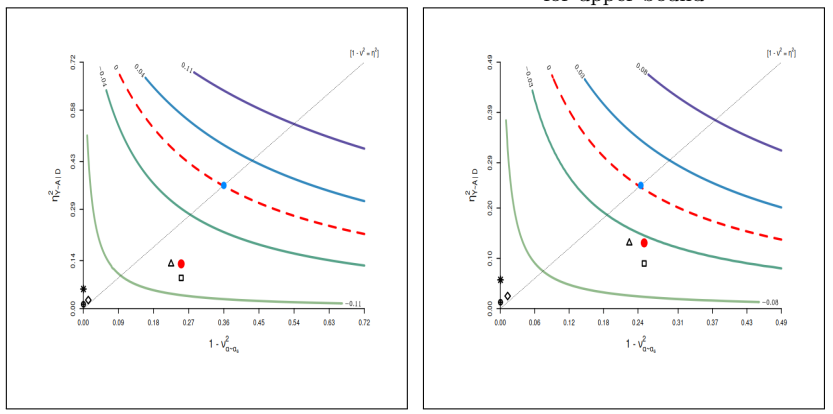


Figure 36: **Wife has no voice in household**

A. Upper bound

B. Upper 95 % confidence bound for upper bound

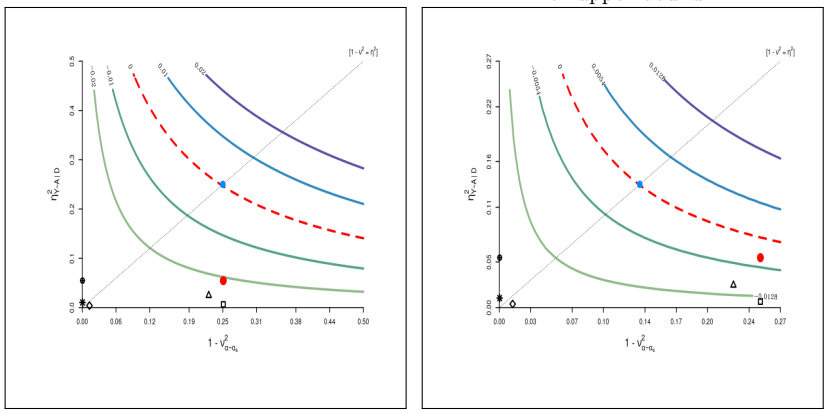
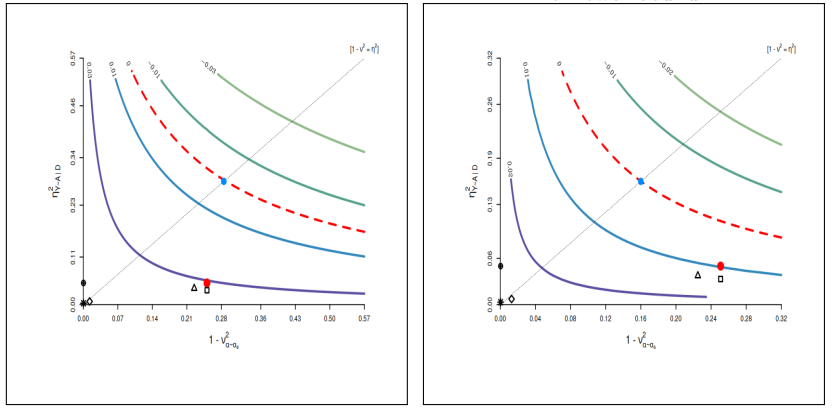


Figure 37: **Report domestic violence acts witnessed by him**

A. Lower bound

B. Lower 95 % confidence bound for lower bound



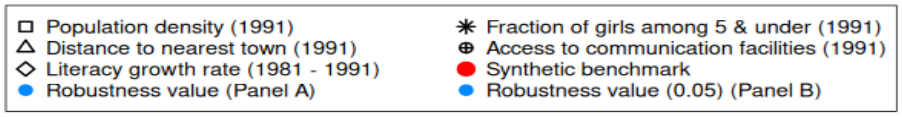
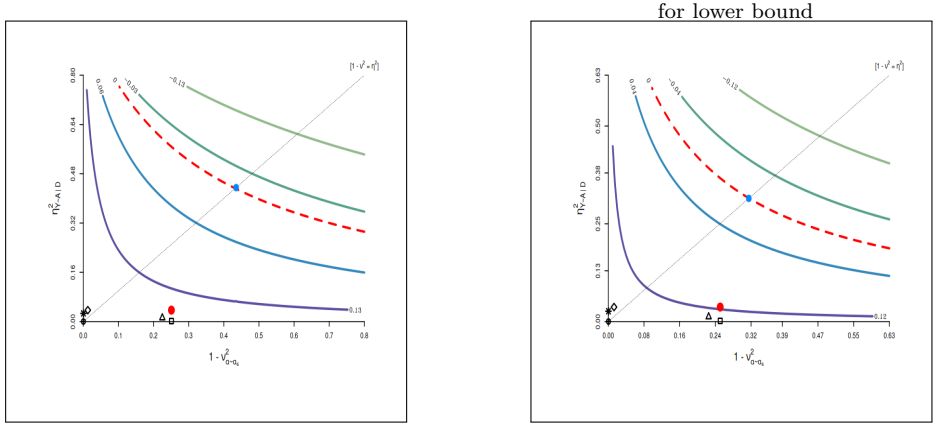


Figure 38: **Demolish illegal liquor shops in the community**

A. Lower bound

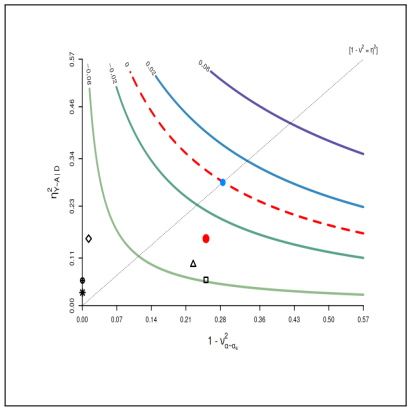
B. Lower 95 % confidence bound for lower bound



Contour plots based on Chernozhukov et al. (2024) based on IPWRA (wife's responses)

Figure 39: Physical abuse of wife

A. Upper bound



B. Upper 95 % confidence bound for upper bound

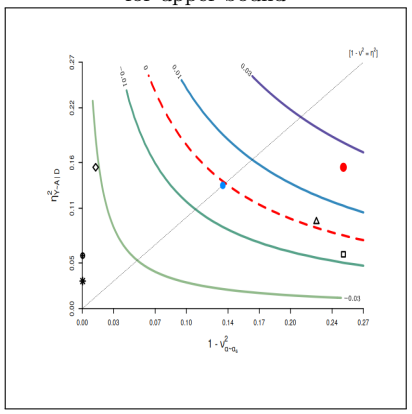
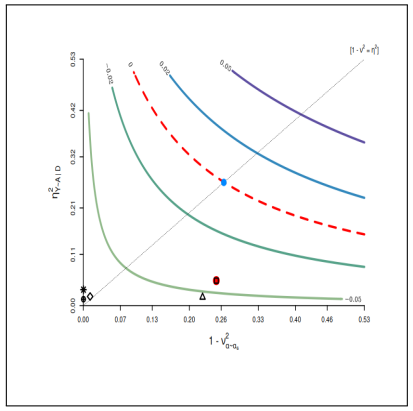


Figure 40: Sexual abuse of wife

A. Upper bound



B. Upper 95 % confidence bound for upper bound

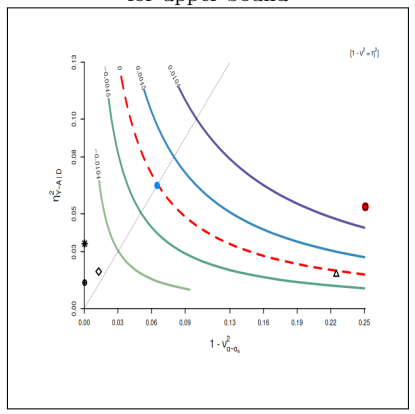
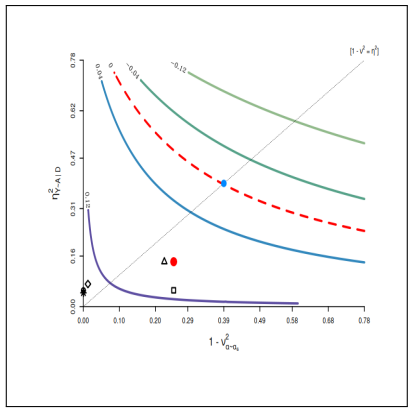
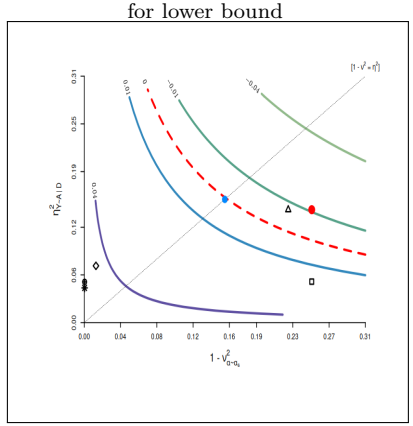


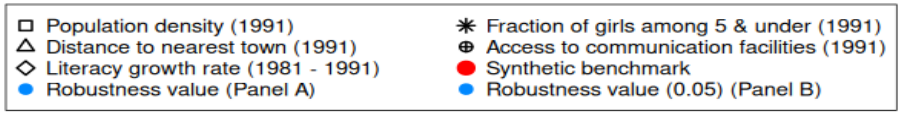
Figure 41: Abuse-free relationship

A. Lower bound



B. Lower 95 % confidence bound for lower bound

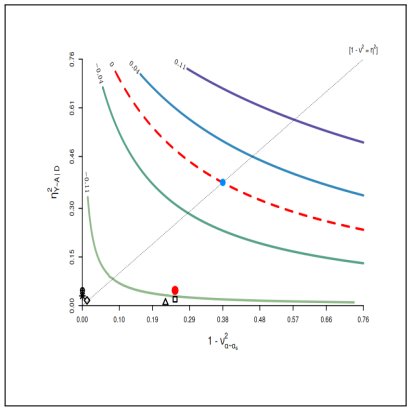




Contour plots based on Chernozhukov et al. (2024) based on IPWRA (wife's responses)

Figure 42: **Wife does not know about the PDVA**

A. Upper bound



B. Upper 95 % confidence bound for upper bound

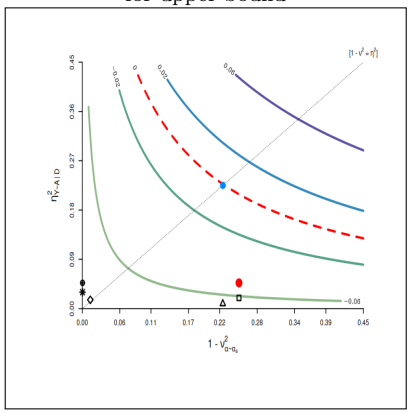
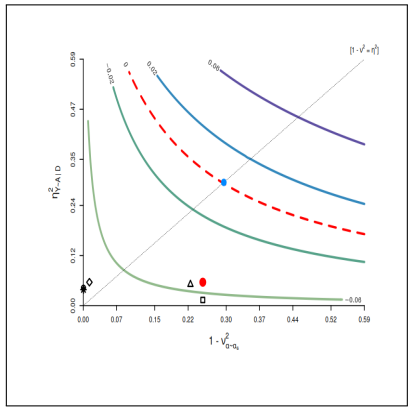


Figure 43: **Wife has no voice in household**

A. Upper bound



B. Upper 95 % confidence bound for upper bound

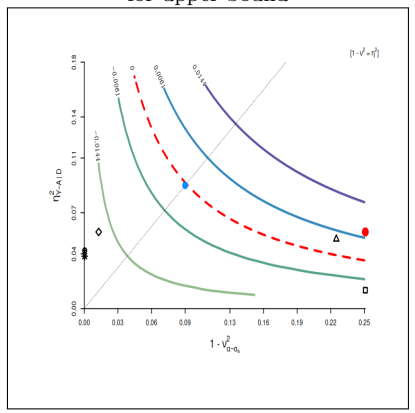
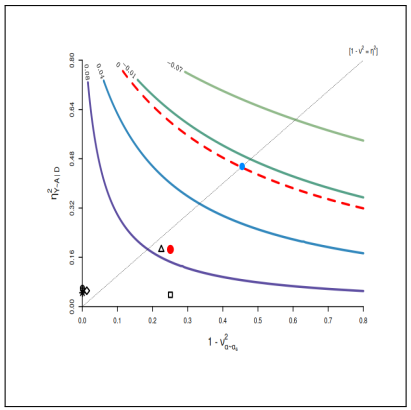


Figure 44: **Report domestic violence acts witnessed by her**

A. Lower bound



B. Lower 95 % confidence bound for lower bound

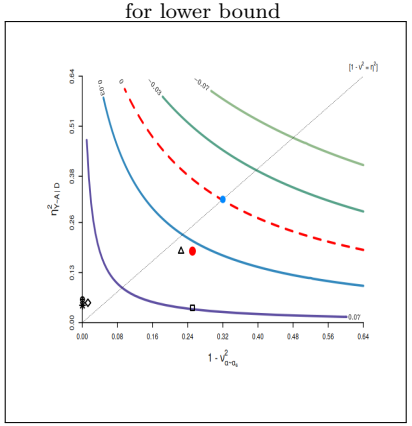
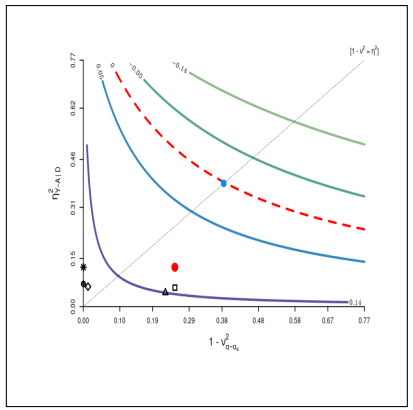
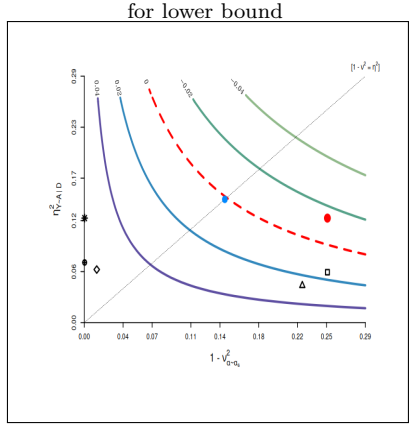


Figure 45: **Demolish illegal liquor shops in the community**

A. Lower bound



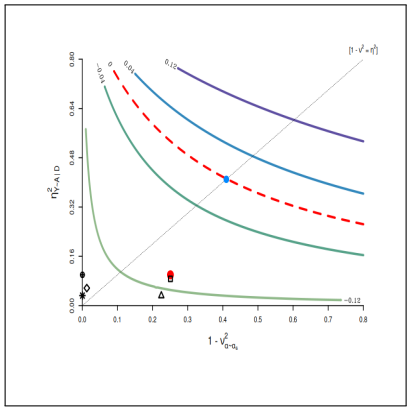
B. Lower 95 % confidence bound for lower bound



Contour plots based on Chernozhukov et al. (2024) based on IPWRA (husband's responses)

Figure 46: Wife beating by husband is justified

A. Upper bound



B. Upper 95 % confidence bound for upper bound

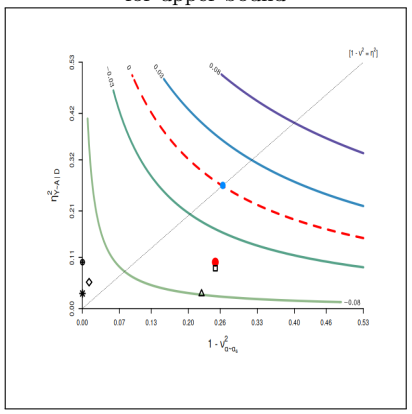
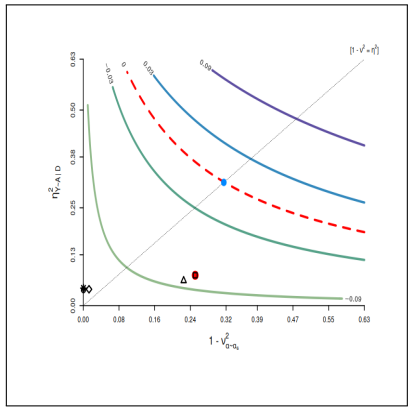


Figure 47: Husband does not know about the PDVA

A. Upper bound



B. Upper 95 % confidence bound for upper bound

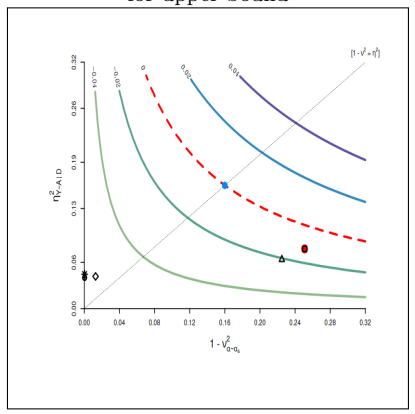
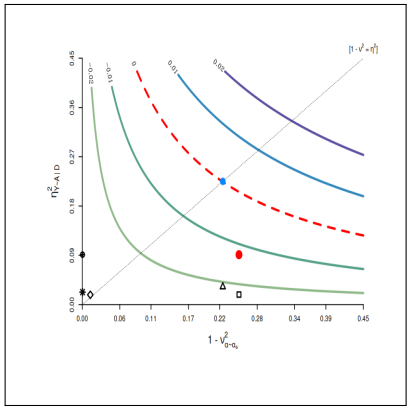


Figure 48: Wife has no voice in household

A. Upper bound



B. Upper 95 % confidence bound for upper bound

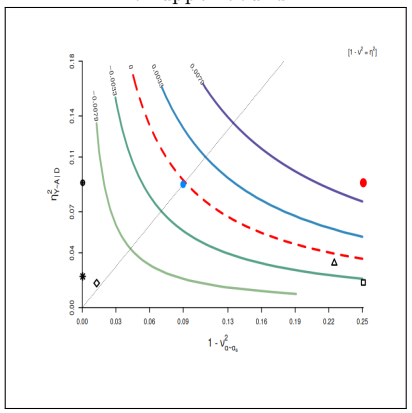
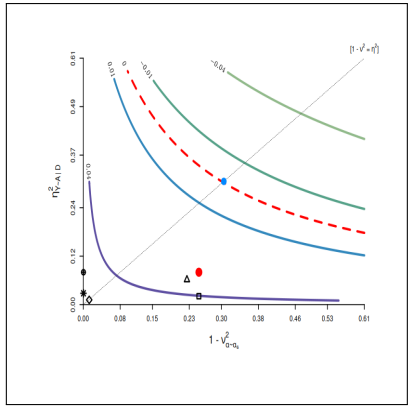


Figure 49: Report domestic violence acts witnessed by him

A. Lower bound



B. Lower 95 % confidence bound for lower bound

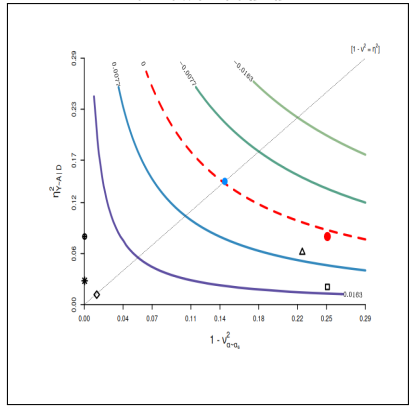
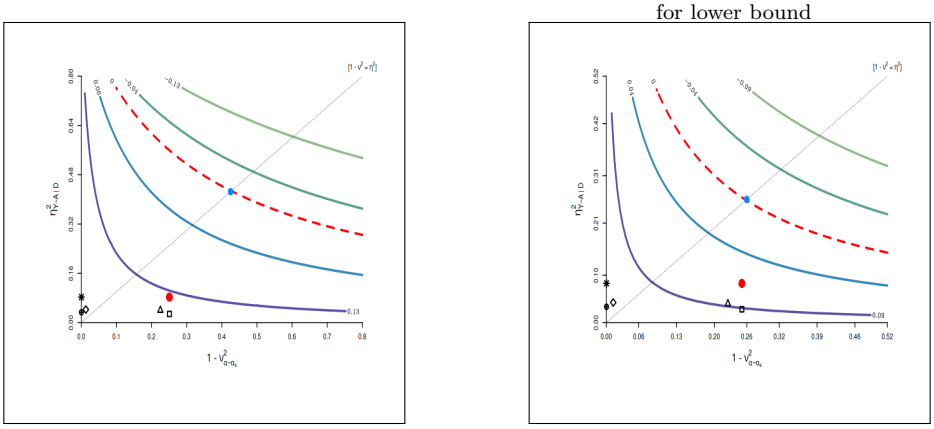


Figure 50: **Demolish illegal liquor shops in the community**

A. Lower bound

B. Lower 95 % confidence bound for lower bound



B Synopsis of Three Jana Sanskriti plays

Shonar Meye (Golden Girl)

Shonar Meye addresses the plight of young women in India. It portrays the life of a girl from childhood to adulthood. The writing of Shonar Meye was a result of 5 years of intensive work in remote areas of the Ganges Delta in West Bengal. At this time *JS* conducted theater workshops with many groups of villagers. The play Shonar Meye is based on the discussions at these workshops, in which individuals shared their experiences of oppression and their daily challenges. Though written almost 15 years ago, it remains an extremely relevant play at the time of this writing.

In the play, Ram babu is a middle-class villager who lives with his wife, son, and daughter. Ram babu favors his son over his daughter because he believes that his son will look after him and his wife in their old age but his daughter will get married and leave for her in-laws' home. His daughter wants to study, but her family will not support her education. The family makes her spend most of her time on daily household chores. Before marriage, prospective in-laws inspect the girl to check whether she is physically suitable to marry their son. The girl passes the inspection, and Ram babu arranges for his daughter to marry the handsome son of a well-to-do family. The groom's family demand a dowry of 10,000 rupees and 110 grams of gold. They ask Ram babu to arrange the dowry by the time of the marriage. Ram babu decides to sell off his land and take a bank loan for the marriage, but fails to pay the dowry by the time of the marriage. The groom's father threatens Ram babu and his family that they will not be able to see their daughter again until he satisfies all their demands. The daughter faces the wrath of her husband's parents because of her father's inability to meet the dowry demands. She has to work very hard. If she makes even a small mistake, she is beaten. The play ends when the daughter confronts her oppressors.

Ekti Meyer Kahini (Story of a Girl)

This is another play that depicts the different stages in the lives of women: the period before marriage, the arrangement of the marriage, and life after marriage. The first part of the play highlights gender inequality. The protagonist, Sankari, is a teenage daughter of a poor agricultural worker. She has an elder brother who is married. Sankari wants to study, but because of her family's poverty she is unable to do so. Her brother and sister-in-law want her to get married and leave for her in-laws' house as soon as possible. The second part of the play showcases dowry-related problems and the lack of choice that Sankari has about when and whom to marry. Sankari's father pays a hefty dowry at the time of her marriage. The last part of the play focuses on the ill effects of early marriage on young girls (13-14 years old) and how the central characters of the family into which a young girl marries (the mother-in-law and husband) become tools of oppression. As punishment for mistakes she made doing household chores, Sankari is sent back to her natal family to bring money to meet the medical expenses incurred by her in-laws when she fell sick. At her father's house, the situation is no better for Sankari: her brother and sister-in-law harass her and put pressure on her father and on her to go back to her in-laws' house. Sankari knows that if she returns to her husband's home without the money her in- laws demand, she will be killed. The play ends as she sees her dilemma – whether to return to her husband's home or stay at her father's home and try and earn a living for herself.

Hay Re Mod (The Curse of Alcohol)

In early 2005, JS organized a sit-in protest against the illegal production and sale of liquor in the region in West Bengal where the teams perform. Villagers blocked the main highway that connects Kolkata to the Ganges Delta. This was the start of an anti-liquor campaign. There is a strong nexus between politicians, illegal liquor shop owners, local government officials, and the police. Prasad Sarkar, one of the protesters, explained the cause of the protest in his indictment of the police: "You are spineless policemen. You find our work illegal [i.e.,

blocking the highway], and you don't notice (chokhe pore na) the illegal production of liquor because it is in your self-interest" (Da Costa 2010). Women and adolescent children bear the brunt of the consequences of alcohol abuse in increased domestic violence and children forced to drop out of school due to the shortage of funds. The play Hay Re Mod (The Curse of Alcohol) presents these problems in the narrative of a woman named Naina. She has two school-going sons and two married brothers-in-law. Naina's husband is a drunkard. He spends all his earnings on alcohol and contributes nothing to run the household or to buy books and school supplies for their sons. From time to time, she has to borrow from her neighbors and do odd jobs outside her home to support the household. When Naina asks him for money, he becomes violent and mercilessly beats her. Naina complains to the head of the village government (the Panchayat). She even goes to the police station to register a complaint. Naina laments that the police "take bribes behind the scenes (pechon theke ghoosh), that is why they cannot find a solution to our problems" (Da Costa 2010). The husband's brothers, too, plead with him to stop drinking, but their efforts fail. The play ends with the brothers moving out of the house and Naina taking her husband to a barely functional rural hospital to get him treated for nausea, abdominal pain, and dizziness – all effects of consuming illicit liquor.

Reference

Da Costa, Dia. 2010. "Have they Disabled Us? Liquor Production and Grammars of Material Distress" in *Development Dramas. Reimagining Rural Political Action in Eastern India.* Routledge: Delhi.

C Sampling Procedure

To select a sample of married women between 18 and 49 years, we use stratified random sampling:

- From each of the 3 control blocks and 3 treatment blocks, we randomly sampled between 1 and 7 Gram Panchayats (GPs).
- From each control GP, we sampled census villages with probabilities proportional to the 1991 population.
- From each treated GP, we randomly sampled census villages from a list provided by JS of the villages where it regularly performs.
- From each census village, we randomly sampled either one or two polling booths from the 2014 electoral list.
- From each electoral list for each selected polling booth, we randomly sampled, in the control villages, 15-35 households and, in the treatment villages, 20 households from the active area (electoral booth where plays were performed) and 15 households from the inactive area (where plays were not performed). We used the electoral list because a voting card is a proof of identity held by most residents of at least 18 years of age the minimum voting age in India).
- The field workers were charged with the task of interviewing one married couple in each household. Each investigator was given details (name, gender, age, and the household head's or husband's name) of the members of households to be interviewed and a list of possible replacement households if no eligible married woman was present and willing to participate in the survey. With the help of a family member, on arrival at the household's home, investigators determined whether an eligible married woman was present. If more than one such woman was present, the investigator-team randomly chose one. In only rare cases did an eligible household member refuse to cooperate.

- The team sought to interview the husband of the selected female respondent wherever he may have been at the time of the wife's interview. The team achieved this in 99 percent of the cases. For the remaining one percent of the married women, another married man of the same or neighboring household was interviewed. The tables in this paper report data only for married couples.

D Survey Questions and Instructions to the Field Investigators

Questions on abuse, asked to the female respondents only

I am going to ask you some questions about your relationship with your husband. Does/did your husband ever:		
	Yes	No
1. *Emotional abuse*:		
a. Say or do something to humiliate you in front of others?	1	2
b. Threaten to hurt or harm you or someone close to you?	1	2
c. Insult you or make you feel bad about yourself?	1	2
2. Physical abuse:		
a.Throw something at you? Twist your arm or pull your hair? Slap you?	1	2
b. Punch you with his fist or something that could hurt you? Try to choke you or burn you? Threaten or attack you with a knife or a gun or any other weapon?	1	2
3. Sexual abuse:		
a. Physically force you to have sexual intercourse with him even when you did not want to?	1	2

Instructions to field investigators on how to ask these questions:

1. a. Suppose your husband does not like the food that you have cooked or he is unhappy about something that you have done. Say, when you go outside the house to fetch water, someone from your parents" house calls you. You receive the call and talk to the person for a few minutes. Observing this, does your husband get annoyed? Does your husband scold you publicly in the presence of your friends and neighbors?

   b. Does your husband threaten you or somebody close to you? Suppose you have burnt the food because you were attending to your child who was crying. Does your husband threaten you by saying that if you repeat this (burn the food), I will accost your brother when I meet him in the market and physically harm him?

   c. Does your husband tell you that you are useless, illiterate, and ignorant? Does he compare you with your neighbor's wife, saying that not only is she beautiful, but she

also runs her household efficiently, helps her children with their homework, and earns money by making hand-made dolls during her spare time?

2. a. & b. Does your husband physically abuse you? Suppose your husband returns home after working in the fields, takes a shower, and sits down for his afternoon meal. The food is not yet ready. He throws whatever he finds within his reach at you. If the meal is too hot, he pulls your hair and slaps you. This is one instance when the husband physically abuses his wife. The husband could also be frustrated that his wife did not bring adequate dowry during their marriage. This frustration is reignited when his friends discuss the huge amount of dowry that another friend of theirs got. Does your husband take out his frustration by physically abusing you? Does the physical abuse take extremely violent forms like your husband strangling you or burning you by pushing you towards the open flame?

3. (The field investigators are trained to modulate their voice to a very low pitch and whisper) Sister, I am going to ask you a personal question concerning your sexual relationship with your husband. Please do not take it otherwise. We are only trying to assess whether you face such problems or have faced them in the past. Suppose one day, you are feeling feverish or exhausted because your child has been crying the whole day and you had a few relatives of your husband over for a meal. At night, your husband may want to indulge in sexual activities with you, but you are not enthusiastic about it. You express your feelings to your husband. Is it the case that in spite of your unwillingness to indulge in sexual activities, you husband tries to force you?

Questions on decision making in the household

(No special instructions to the field investigators except for reading out the options)

I am going to ask you about who makes the decisions in the following situations					
Who decides about:	Respondent	Husband (wife)	Both	Family	Other
a. Children's education	1	2	3	4	6
b. The family's health care needs	1	2	3	4	6
c. Purchasing major household items	1	2	3	4	6
d. Purchasing minor household items	1	2	3	4	6
e. Visits to the wife's family or relatives	1	2	3	4	6
f. Marriage of your children	1	2	3	4	6
g. Number of children to have	1	2	3	4	6
h. Use of contraceptives	1	2	3	4	6

Questions on marital control, asked to female respondents only

I am going to ask you about some situations which happen to some women. Please tell me if any of these situations arise in your relationship with your husband.			
	Yes	No	DK
a. He (does/did) not trust you with any money	1	2	9
b. He does not permit you to meet with your female friends	1	2	9
c. He tries to limit contact with your natal family	1	2	9

Instructions to the field investigators on how to ask these questions:

a. Do you have any money you can use in times of emergency when your husband is not at home?

b. Do you socialize with other wives or your female friends in the village? Suppose you have finished all your household chores, and the children are away at school. A close friend of yours who had been your neighbour married someone from another village but was currently visiting her natal family. You decided to pay a short visit to the neighbor's house to meet your friend. In such cases, will your husband get annoyed? Alternatively, if you requested that on the way back after picking up your children from school, you would like to stop at your friend's house for some time, would he grant your request?

c. Does your husband get annoyed if someone calls you daily from your parents' house or if you call your parents frequently? Does he prevent you from maintaining a close relationship with members of your natal family?

Questions on attitudes towards wife beating

In your opinion, is a husband justified in hitting or beating his wife in the following situations?			
	Yes	No	DK
a. If she goes out without telling him	1	2	9
b. If she neglects the house or the children	1	2	9
c. If she argues with him	1	2	9
d. If she refuses to have sex with him	1	2	9
e. If she shows disrespect to her in-laws	1	2	9
f. If he suspects her of being unfaithful	1	2	9
g. If she is does not have a male child	1	2	9

Instructions given to the field investigators on how to ask these questions:

a. Suppose while her husband is at work, a woman finds that she is missing an essential ingredient for cooking that is available in the neighborhood store. Or suppose that she gets news that her mother, who lives in a neighboring village, has fallen down and sprained her ankle. She goes to the store / visits her parental home. When her husband returns home, she informs him that she has gone to the store/visited her natal home. Suppose that, on hearing this, the husband angrily questions her: Why did she go out of the house without asking him? Why hadn't she arranged in advance to have all essential ingredients at home? Since her father and brother are living in her natal home, why did she have to rush to attend to her mother? If the annoyance expressed by the husband leads to him to assault his wife, is he justified?

b. A couple have 3 (replace with the correct number) children. There is a lot of housework that the wife needs to complete before her husband returns from the fields at lunchtime. The wife has bathed her children and dressed them in fresh clothes. After some time, she finds them playing in the dirt. While the wife is finishing her cooking, she finds

that her infant child has started to cry; or that her children are quarreling and crying. Her husband comes home and blames her. He accuses her of neglecting the household and the children. If the annoyance expressed by the husband leads him to assault his wife, is he justified?

c. Suppose a wife fails to serve a meal on time because she had to prepare her children for school, clean the house, and complete the cooking. Her husband returns home from work (in the fields) and finds that his wife has not finished cooking. The wife tries to explain to her husband why she had been delayed in cooking the meal on that day. But he does not listen and accuses her of neglecting her household chores. He blames her parents for not teaching her how to run a household efficiently etc. This leads to an argument between the two and finally the husband assaults his wife. Is this action by the husband justified? (voice ofthe investigator should be very soft and sympathetic while asking this question)

d. Suppose one day, a wife is feeling feverish or exhausted because her child has been crying the whole day and she also had to serve a lavish meal to a few relatives of her husband. At night, her husband may want to indulge in sexual activities with his wife, but she denies him. She explains her reasons to her husband. If the husband gets angry with his wife for denying him his rights and assaults her, is he justified?

e. Suppose a woman's child is crying and her mother-in-law accuses her that she does not take care of her child properly, does not feed the child on time, etc. Due to incessant crying of the child, the wife's frustration builds up and she talks back to her mother-in-law. The mother-in-law takes this as an insult and complains to her son when he returns home during lunchtime. He gets angry with his wife for insulting his mother and physically assaults her. Is he justified in his actions?

f. Suppose a wife meets a male relative/acquaintance from her natal village when going to the market (or suppose she discusses with her child's male teacher her child inatten-

tiveness towards studies). Whenever her husband sees his wife talking with another man, he grows suspicious that she is cheating on him. Because of his suspicions, he scolds her, verbally abuses her, and finally resorts to physical violence. Do you think the husband is justified in his actions?

g. Suppose a couple already has 3 children (replace with the correct number of children). The wife does not want any more children because she realizes that their current household financial situation will not allow them to provide the additional child with adequate nutrition and education. But her husband wants another child because they only have daughters. He wants a son to carry his lineage. The wife tries to reason with her husband. This leads to a verbal argument between the two. Finally, the husband assaults his wife in anger. Do you think the husband is justified in his actions?

Questions on community actions

No special instructions given to the field investigators when asking these questions.

Please tell me if you would willingly participate in the actions described below			
	Yes	No	DK
a. Participate in demolishing an illegal liquor shop in your neighborhood/village?	1	2	9
b. Report to anyone an act of domestic violence that you have witnessed?	1	2	9

Questions on awareness about laws against domestic violence

No special instructions were given to the field investigators when asking these questions.

Please tell me if you know about the following			
	Yes	No	DK
a. Do you know that a woman can file a complaint in the police station against any male perpetrator or any female or male relatives of the husband who has perpetrated domestic violence against her	1	2	9
b. Do you know that a woman who is a victim of domestic violence can register a case if she is restrained from entering any portion of the shared household in which she resides	1	2	9
c. Do you know that a woman who is a victim of domestic violence can claim monetary compensation for medical treatment, loss of earnings, and maintenance for herself and for her children from her husband?	1	2	9

Questions on spousal abuse due to drinking or substance abuse asked to the female respondent

Do any of these situations arise in your relationship with your husband?			
	Yes	No	DK
a. Does your husband drink alcohol or use drugs?	1	2	9
b. Does your husband verbally or physically abuse you while intoxicated?	1	2	9

Instructions to the field investigators on how to ask these questions:

a. Does your husband consume any intoxicant, like alcohol or drugs?

b. (Ask after confirming with the respondent that her husband consumes alcohol or takes drugs.) When your husband comes home intoxicated, does he pick a quarrel with you? Does he verbally or physically abuse you?

Question on spousal abuse due to drinking or substance abuse asked to the male respondents

Does any such a situation arise in your relationship with your wife:			
	Yes	No	DK
Do you verbally or physically abuse your wife when you are intoxicated?	1	2	9

Instructions given to the field investigators on how to ask this question:

(Ask after confirming with the respondent that he consumes alcohol or takes drugs) 'When you come home intoxicated, do you sometimes pick a quarrel with your wife? Do you verbally or physically abuse her?