Tackling Sexual Harassment: Short and Long run Experimental Evidence from India*

Karmini Sharma†

September 20, 2022
Click here for the latest version

Abstract

Sexual harassment imposes substantial economic costs on the victims yet, there is limited evidence on how to effectively deter it. I show that men can be successfully engaged to reduce sexual harassment using an RCT targetting sexual harassment awareness training at men and women separately in colleges in New Delhi. I show that the training with men was able to reduce overall sexual harassment by 0.06 sd with a 100% decrease in extreme forms of harassment by reported by women in their peer groups. However, the training also reduces inter-personal relationships between men and women. I find that this is driven by women's choices, using a lab-in-the-field experiment in which women prefer to cooperate with women rather than men on an experimental task. Using a signalling framework, I show that this is consistent with some men undertaking "good" behaviours even though they would prefer to harass women, to avoid disapproval from their peers. I find that there is an increase in men's perception of peer disapproval against sexual harassment and no change in their intrinsic attitudes towards it, consistent with the theoretical predictions. Long term surveys over one to two years after the intervention show that the effects on sexual harassment become weaker while on opposite sex relationships are persistent. I cannot reject a null effect on sexual harassment and opposite sex relationships of a similar intervention that was delivered exclusively to women in a separate college.

JEL codes: D91, J16, J28, K42, O12.
Keywords: Gender, sexual harassment, deterrence, beliefs, relationships.

*I am grateful to James Fenske, Clément Imbert and Christopher Roth for their guidance and support for the project. Funding for this project was provided by the UK Foreign, Commonwealth Development Office, awarded through the Crime and Violence Initiative at J-PAL. I am grateful for all the support from JPAL-SA staff, IWWAGE-IFMR, Safecity for their grant support for a part of the project. I thank Prakriti Singh and Ayushmita Samal for their excellent research assistance. I am indebted to Pooja, Najma, Rahul and Nazmeen for field work. I thank Vanshika, Siddharth and Kajal for excellent delivery of trainings. I thank Debraj Ray for his guidance on the theoretical framework of the paper. I thank Zeinab Aboutalebi, Arun Advani, Robert Akerlof, Anik Ashraf, Haseeb Ashraf, Lori Beaman, Sonia Bhalotra, Apurav Yash Bhatiya, Girija Borker, Stefano Caria, Costas Cavounidis, Olle Folke, Felix Forster, Maitreesh Ghatak, Bishnu Gupta, Sinem Hidir, Seema Jayachandran, Gabriel Kreindler, Anandi Mani, Sharun Mukand, Dev Patel, Gautam Rao, Johanna Rickne, Andreas Stegmann, Teodora Tsankova, and Todor Tochev for their feedback. This project received IRB approval from Institute for Financial Management and Research and University of Warwick and pre-registered at the AEA registry with id number AEARCTR-0003678. All errors are my own.

†Stanford University, karmini3[at]stanford[dot]edu

1 Introduction

Sexual harassment is a ubiquitous phenomenon for women.[1] For instance, more than 1 in 2 women in the EU and 4 in 5 women in Delhi, experience sexual harassment over their lifetime.[2] Pervasive sexual harassment can have a debilitating impact on the psychological, economic, and social lives of women (Jayachandran (2015); Borker (2017); Talboys et al. (2017); Folke and Rickne (2020); Azmat et al. (2020)). Potential victims of sexual harassment were more likely to quit their workplace in Sweden and choose lower quality educational institutions in Delhi to avoid sexual harassment (Borker (2017); Folke and Rickne (2020)). Thus, it is crucial to understand what works to deter sexual harassment. Sexual harassment awareness training has long been advocated by policymakers and academics for tackling sexual harassment in universities and workplaces.[3] However, there is a lack of causal evidence on its impact on sexual harassment incidence and the mechanisms behind it.[4] Such training, when provided with one's peers, can affect an individual's own attitude towards sexual harassment but also his or her perception of which behaviours are 'approved' by their peers. This can affect other interactions between men and women like romantic relationships, friendships, and their professional networking with each other.[5] Understanding the effects of this training on such relationships is crucial for organizations where social incentives matter for performance. In this paper, I provide experimental evidence on the effect of sexual harassment awareness training on sexual harassment incidence for women, in a sample of college students. I also study its impact on relationships between men and women (henceforth, opposite sex relationships).[6]

I collaborated with the NGO, Safecity, to randomly provide this training to only male students in two colleges (henceforth, the male intervention), giving a sample of nearly 3100 students. This training was provided by NGO trainers to all men in randomly selected classes for a total of three to five hours per class. The training had two main components: awareness and empathy building. The first component provided men with information on sexual harassment, the laws against it, and simple ways to help a survivor of sexual harassment. The empathy building component was added to minimize the chances of backlash from men that has been documented before (Bingham and Scherer (2001)). This was mainly done through a discussion of anonymous narratives that I collected from women in men's colleges. These

[1] Sexual harassment is defined as any unwanted and unwelcome behaviour of a sexual nature (Sexual harassment of Women at Workplace Act, 2013 India)).

[2] Safe Cities Free of Violence Against Women and Girls Initiative: Report of the Baseline Survey Delhi (UN women, 2010) and Violence against Women: An EU wide survey (2015) by FRA.

[3] This awareness training has been advocated by lawmakers *SHWA* 2013, *EEOC* for US and academics (Fitzgerald and Shullman, 1993). This training is mandatory in many countries in Europe, Asia and 21 states in the US and recommended by majority of the countries.

[4] Roehling and Huang (2018) provides a review.

[5] See Zhu (2019) for importance of networks developed in college; Beaman and Magruder (2012) and Sacerdote (2001) on utility of such relationships in the labour market.

[6] A college can have between 2000 to 5000 enrolled students in a given academic year and has approximately 30 courses (degree programs or majors) available for students to enrol in.

narratives were discussed during the training to help men understand women's perspective within a sexual harassment incident. Both these components of the training were simple and basic, which makes them replicable in other institutional settings as well.

To ensure that women's understanding of sexual harassment between treatment and control classes was the same, I provided all women in all classes with information on sexual harassment at the baseline. I randomized the provision of this same information for women in a separate college (henceforth, the female intervention). I randomized at the class level, the proportion of women who received this information. The follow-up survey for both the interventions was done three to four months after the treatment.

I find that the male intervention leads to a significant fall in overall sexual harassment by 0.06 sd. In particular, I find a significant fall in extreme forms of sexual harassment by 1.1 p.p. as reported by women from within the same classes. This translates to 51 fewer women out of 1200 experiencing extreme forms of sexual harassment over one academic year. Extreme forms of sexual harassment include groping, pinching and other extreme acts without consent of the victim. To rule out that these results are driven by changes in reporting behaviour of women, I show that the treatment had no such negative effect on women's reporting of sexual harassment from men outside college or from men in different classes. I also undertook various steps to minimize bias from under-reporting, stigma, selection into the sample, under-detection and privacy issues in collecting sexual harassment data due to sensitivity of the topic (Aguilar et al. (2020)).[7] Next, I find that there was a significant negative effect of the treatment on an opposite sex relationships index, measuring different forms of relationships between men and women, by 0.13 sd. In particular, there was a decline in romantic relationships by 1.3 p.p., suggesting higher gender-segregation.

To understand why the male intervention leads to a fall in sexual harassment and opposite sex relationships, I use a signalling framework adapting Bénabou and Tirole (2006) and Bursztyn et al. (2020a). Since all the men in treatment classes were provided with the training together with their male classmates, two mechanisms could be at play. The training can affect not only trained men's own attitudes, but also their beliefs about their peers' disapproval of sexual harassment. I model interactions between men and women with these two possible mechanisms in mind. The results are consistent with the second mechanism. It leads men, who intrinsically prefer to harass women, to instead undertake 'good' behaviours towards women after the male intervention. This makes women more cautious in forming relationships with them, conditional on even 'good' behaviour. This leads to a fall in sexual harassment (because of men's pooling) and a fall in relationships (because of women's caution).

In particular, I show that the framework yields different predictions for the two possible mechanisms. Both mechanisms lead to a reduction in sexual harassment but only a higher

[7]Reporting of sexual harassment can lead to victim blaming, backlash from the perpetrator and other forms of retaliation (Dahl and Knepper (2021)).

perceived social disapproval from peers can lead to a decrease in opposite sex relationships. A change in intrinsic attitudes can only lead to an increase in such relationships. Social disapproval can be imposed by mutual peers of potential perpetrators and victims (Folke and Rickne (2020)) through social image and reputation costs on perpetrators.[8] Such concerns have been extensively studied in the context of voting, conformity to social norms, donations, and preventive health behaviours.[9]

Several features of my data allow me to test for the two mechanisms. First, my results on primary outcomes, that is, a fall in sexual harassment and opposite sex relationships are consistent with the mechanism of higher perceived peer disapproval. Second, I collected direct survey questions from men and women on their perceptions about others. I find that the male intervention increased men's perceived social disapproval of sexual harassment by 0.056 sd. Men also report that attitudes of others in their class changed significantly after the treatment. Additionally, I show that post-intervention, men believe that their female classmates are more likely to report sexual harassment to their peers rather than to the legal complaints committee of their college. Women, however, did not think that men's attitudes changed in their class after the treatment. All these results are consistent with the second mechanism. Third, I also measured men's attitudes using a list experiment and a volunteership exercise. These two measures help alleviate experimenter demand effect concerns with collecting self-reported attitudes on sexual harassment. I cannot reject a null effect of the intervention on intrinsic attitudes towards sexual harassment (minimum detectable effect, MDE of 0.03). Fourth, I also provide other indirect evidence consistent with the theory. Evidence from a lab-in-the-field experiment with the same students suggests that women's choices were driving the fall in opposite sex relationships, consistent with women's increased caution. In particular, I find that the women in treatment classes preferred to cooperate with women over men in an experimental task by 13 p.p.. I also find that the reduction in opposite sex relationships was driven by women in their first year (freshman) of study, consistent with a lack of information about men's types driving women's behaviour.

To rule out alternative mechanisms, I show that men's perception of legal implications of sexual harassment did not change. I also find that women's relationships declined only with men in their own class, but increased with men from outside the class after the male intervention. Thus, a change in men's behaviour within the treated classes drove women's behaviour rather than women reducing relationships with all men. There can be concerns that sexual harassment decreased only because women's relationships with men decreased due to the treatment. I show that women in control classes, who were single, also reported facing

[8]Throughout the paper, I implicitly refer to men as potential perpetrators and women as potential victims. This corroborates with official reports that show that majority of the perpetrators of sexual violence are men and majority of the victims are women (UN Women, 2015).

[9]Some of these studies include Karing (2018); DellaVigna et al. (2016); Bénabou and Tirole (2006); Bernheim (1994); Macchiavello and Morjaria (2015) and Greif (1989).

sexual harassment. Further, I show that the decline in opposite sex relationships does not mediate the effects on extreme forms of sexual harassment. The treatment did not have any effect on men's choices in the lab-in-the-field experiment, which rules out that men became more cautious in interacting with the women. Further, I show that all the results are robust to multiple hypothesis testing and randomization inference. Finally, within college spillovers can bias my results downwards. Thus, my results should be seen as a lower bound on true effects of the male intervention.

In contrast, I cannot reject null effects of the female intervention on sexual harassment and on opposite sex relationships. I can rule out effect size of 0.1 or more for the female intervention on these two outcomes.

In long run surveys undertaken approximately one to three years after the intervention, I find that women report lower intermediate forms of harassment (stalking, online harassment, among others.), but results on overall sexual harassment are noisier and insignificant. These coefficients are, however, larger than in the short run (-0.5sd versus -0.8 sd in long run). The estimates are likely rendered noisier since I asked women to recall sexual harassment that happened over 6 to 7 months after the intervention. But because of the pandemic, the surveys with women themselves happened much later than the recall period. I find however, women continue to report lower romantic relationships by 47% with men in treatment classes about one to two years after the interventions. This shows that women's beliefs about men likely did not change in the long run and that effects on opposite sex relationships were persistent.

By constraining women's participation in labour markets, human capital accumulation and mental health, sexual harassment impedes women's empowerment (Duflo (2012)). This paper contributes to evidence from developing countries showing that enlisting men for women's empowerment can succeed in developing countries because of greater relative freedom that men might have, to act on their beliefs and/or preferences (Dhar et al. (2018); Bursztyn et al. (2018); Haushofer et al. (2019); Ashraf et al. (2020b)). A number of papers have studied ways to empower women using various skills training programs (Ashraf et al. (2020a); McKelway (2020); Bandiera et al. (2020); Edmonds et al. (2020); Buchmann et al. (2021); Gulesci et al. (2021)) or improving attitudes of communities and families of these women (Abramsky et al. (2014); Dean and Jayachandran (2019); Banerjee et al. (2019); Green et al. (2020)). I show that one other way could be to directly engage men by changing their perception of the social environment.[10]

I add to the literature on endogenous network formation and in particular, gender segregation. I contribute by studying how the training affects opposite sex relationships in colleges. I

[10]There are studies in social psychology that look at the impact of sexual harassment training on attitudes (Roehling and Huang (2018); Antecol and Cobb-Clark (2003); Bingham and Scherer (2001)) but cannot causally track effects on sexual harassment incidence or opposite gender relationships. This paper fills this gap as well.

show whether sexual harassment awareness training can exacerbate or reduce gender segregation. I also use a lab-in-the-field experiment to show whether segregation is a result of men's or women's decisions and thus, am able to highlight the mechanisms behind it. While a huge literature shows how networks affect economic and non-economic outcomes, relatively fewer papers study the impact of randomized interventions on networks. These papers study relationships for informal finance, legislative activity, communication among others (Mayer and Puller (2008); Comola and Prina (2014); Banerjee et al. (2018); Dupas et al. (2019); Canen et al. (2019); Jäckering et al. (2019); Ru and Townsend (2020)). A recent upcoming literature studies effects of gender segregation on attitudes towards gender identity and women's labour market outcomes (Miller et al. (2019); Jayachandran (2020); Dahl et al. (2020)) and causes behind it (Jayachandran (2015); Jayachandran (2020)). Gender segregation can limit women's networks affecting their information acquisition and restrict their labour market opportunities (Field et al. (2010); Field et al. (2016); Beaman and Dillon (2018); Beaman et al. (2018)). Given that sexual harassment awareness training is becoming increasingly common and mandatory, results of my paper are pertinent for environments where collaboration may be important.

This paper is closely related to the literature on violence against women including sexual harassment. I contribute by studying what helps to deter sexual harassment and measure effects on opposite sex relationships. I also highlight the role of peers' disapproval in such settings for deterring undesirable behaviours. Much of the upcoming literature focuses on street harassment and its consequences for women in developing countries (Borker (2017); Kondylis et al. (2019); Aguilar et al. (2020)). Previous studies have focused largely on intimate partner violence studying, for instance, effects of cash transfers, gender wage gaps, female labour force participation or motives of men behind it (Bloch and Rao (2002); Aizer (2010); Anderberg and Rainer (2013); Erten and Keskin (2018); Anderberg et al. (2018); Haushofer et al. (2019); Calvi and Keskar (2020); Kotsadam and Villanger (2020)). Relatively little attention has been paid to sexual harassment in workplaces or educational institutions. Few studies show repercussions of workplace sexual harassment for pay inequality and labour market outcomes for the victimized (Basu (2003); Antecol and Cobb-Clark (2006); Hersch (2011); Hersch (2018); Folke and Rickne (2020)). Lindo et al. (2018) study effect of college partying on sexual assaults on campuses which is a setting similar to mine.

Finally, this paper adds to the literature on social image concerns. I contribute to this literature by showing that social image concerns can be activated through higher awareness and can deter undesirable behaviours. This literature shows that perception of what others think can drive one's behavior whether the perceptions are correct or not (Bénabou and Tirole (2006); DellaVigna et al. (2012); Bursztyn and Jensen (2015); DellaVigna et al. (2016); Bursztyn et al. (2020a); Bursztyn et al. (2020b); Bursztyn and Yang (2021)).

I discuss the context of these colleges in section 2, the theoretical framework in section 3, the details of the intervention in section 4, the experimental design in section 5, the results in section 6 and conclude in section 7.

2 Context

UNDP defines sexual harassment as "any unwelcome sexual advance, request for sexual favour, verbal or physical conduct or gesture of a sexual nature, or any other behaviour of a sexual nature that might reasonably be expected or be perceived to cause offence or humiliation to another person." Laws against sexual harassment have become increasingly common as shown in appendix figure E.3. Indian law identifies sexual harassment as "any unwanted or unwelcome behaviour of a sexual nature" (SHWA, 2013). Although there is no data collected on a global scale on awareness of sexual harassment, there have been surveys on beliefs about prevalence of sexual harassment. These surveys show underestimation of sexual harassment by both men and women accross countries.[11] I show this in beliefs data that I collected at the baseline from one of the colleges in figure E.4. The figure gives the distribution of beliefs of men and women over the entire range of hypothesized prevalence of sexual harassment for women in their class. Most men and women underestimate the prevalence of sexual harassment but men more so than women.

As mentioned, I collaborated with three colleges in one of the Universities in Delhi for the paper. College students in the age group of 18 to 21 are particularly suited for the training due to their willingness to discuss the nuanced and sensitive topic of sexual harassment. More importantly, sexual violence on campuses is pervasive and a key focus of various NGO's and policy makers (RAINN).[12]

To put the collaborating colleges in context, I present the rate of prevalence of sexual harassment categorized by intensity. I adapt the sexual harassment experiences questionnaire (SEQ henceforth) that is widely used to measure sexual harassment for colleges and work-places in social psychology (Fitzgerald et al. (1995); Fitzgerald (1988)). I asked women at the baseline about their exposure to different types of sexual harassment incidents, in the two months preceding the survey.[13] These incidents could be of low, intermediate or extreme intensity. Mild events include sexual remarks, jokes, being repeatedly asked out on a date and intermediate events include physical intimidation, stalking, staring, online sexual harassment. Extreme events pertain to physical acts of fondling, groping or sexual assault. Such classification has been previously used by U.S Merit System Protection Board (USMSPB, 1981, 1987). The summary is shown in table 1. Low and intermediate intensity events were highly

[11] See Figure E.2 in the appendix
[12] https://www.rainn.org/statistics/campus-sexual-violence
[13] Elaborated in appendix A

common at 44 to 47%, respectively. Prevalence rate of extreme events was also high with 16% women reporting exposure to such events. Given that this was a recall for the preceding two months, this is a high prevalence of sexual harassment.

Next, I show who perpetrated these sexual harassment incidents in figure 1. Around 12% of sexual harassment incidents of low intensity type, 8% of medium intensity type and 3% of extreme intensity type of sexual harassment events were perpetrated by someone from within the college over the period of two months. The majority of the perpetration came from someone outside college. However, perpetration from inside college comes from someone the female students knew or would most likely come in contact with on a repeated basis. This has the potential to be more harmful. A caveat of this data is that I cannot parse out differences in number of incidents from overall exposure to sexual harassment. Given that sexual harassment from someone inside the college is repeated, number of incidents would be much higher. Given the discussion above, I would interpret the effects on sexual harassment as a lower bound on the effects on number of incidents atleast in absolute terms.

Finally I provide where my collaborating colleges stand in comparison to other colleges in the same university in figure E.1. In particular, these colleges fall in the middle of college quality distribution as measured by score cut-offs used to admit students. This also improves the external validity of the experiment.

3 Theoretical framework

To understand the treatment effects of the interventions better, I adapt the framework commonly used for studying social image incentives behind different decisions (Bénabou and Tirole (2006); Bursztyn et al. (2020a)).

3.1 Social Environment

I set up a signalling framework, where men, M, are senders of signals and women, W, are receivers of those signals. M and W are paired randomly with each other. I assume that men can take two types of actions towards women, b and g. b are sexually harassing behaviours and g are non-sexually harassing behaviours. M can be one of the two types; bad (B) or good (G). W, who are the receivers of M's actions, decide whether to accept their actions or not. A relationship (romantic or friendship) is formed only when W accepts M's actions. However, M is still able to sexually harass W even if she does not accept his action. The key idea in the model is that women exercise some degree of control on which type of man they form a relationship with. They aim to avoid B type men so as to prevent future abuse and harassment within a relationship. A proportion p of men are of B type. The action space for M is $a_i \in \{b, g\}$ and for W is $a_w \in \{Accept, Reject\}$. I assume a presence of

observers (classmates) who can approve or disapprove of men's types on the basis of their actions which are assumed to be observable.[14] They can impose social disapproval costs, D, on those perceived to be B types.

Both types of M get 0 utility from undertaking b. But the G types receive a positive intrinsic utility k from doing g while the B types suffer a psychic cost c_i if they undertake g where $c_i \sim f(.)$ over $[0, \infty)$. Thus, a B type man is characterized by (t_i, c_i) where t_i is the broader type, B while a G type man only has broad dimension. Women form same beliefs as the social environment (classmates) conditional on actions of men which I will depict by P(.). Men's utility is characterized by:

$$U(t_i, a_i) = \underbrace{I(W \text{ accepts } a_i)}_{\textit{Pairing} \text{ utility}} - \underbrace{c_i I(a_i = g, t_i = B)}_{\textit{Psychic} \text{ costs for B types}} + \underbrace{kI(a_i = g, t_i = G)}_{\textit{Intrinsic} \text{ utility for G types}} - \underbrace{DP(t_i = B|a_i)}_{\textit{Social} \text{ disapproval}}$$

$I(.)$ is an indicator function that takes a value 1 if the event is true. The first term gives the utility from forming a relationship with a woman (normalized to 1), the second term depicts psychic cost incurred if B type man has to do g, the third term is the intrinsic utility that G type gets from undertaking g and the last term depicts the social disapproval that M suffers to the extent that he is perceived to be a bad type man. If a woman accepts an action from a man, she receives u if $t_i = G$ and $v(D)$ if $t_i = B$ and 0, if she rejects. I assume that $u > 0 > v(D)$. A woman's dis-utility from being matched with a B type man is dependent on D. $v(.)$ is assumed to be continuous and differentiable function of D. I assume this because an increase in disapproval against B type men also decreases a woman's costs from being with a B type man (like being blamed for sexual harassment if she reports him, costs of reporting a B type man once she realises his type among other forms of support.). Thus, I assume $v'(D) \geq 0$. A woman never accepts a man if he undertakes b because that is legally sexual harassment and we assume she is aware of this.[15] Thus the only way a man can match with her is through g. Recall, that the social environment also holds same beliefs as the woman.

3.2 Timing

The timing of the game is as follows:

1. Nature chooses type of M given the probability p with which M is of B type.

[14] Even if not observable, I assume that women who are receivers of those actions can tell their peers about actions taken by men towards them. I find empirically that women were more likely to report to their peers about a sexual harassment incident after treatment.

[15] This assumption can be justified because all women in all classes were provided with information on sexual harassment in the baseline.

2. M takes action a_i towards W: $a_i \in \{b, g\}$.

3. W observes M's actions and updates her beliefs about M's type: $Pr(t_i = B|a_i)$ and $Pr(t_i = G|a_i)$.

4. W decides whether she will accept or reject his actions: $a_w \in \{Accept, Reject\}$.

5. A relationship is formed if W accepts a_i and not otherwise. The game ends after this.

3.3 Equilibrium

Given woman's utility, it is easy to see that she will follow a cutoff strategy. Conditional on any action a_i, she will *Accept* iff $P(t_i = G|a_i) \geq \frac{-v(D)}{u-v(D)}$ and reject otherwise. I focus on only partial pooling equilibrium where women follow a mixed strategy, that is, where they are indifferent between *Accept* or *Reject*.

Definition 1. *A partial pooling equilibrium of this game is characterized by:*

1. *The equilibrium strategy of the two types of men: For G type man* ($a_G \in \{b,\ g\}$) *and for B type man* ($a_B \in c$ *where* $c \in [0, \infty)$).

2. *Beliefs of the social environment and the woman which are given by* $P(c^*)$ *where* $P(c^*) = Pr(t_i = B|a_i)$ *and* $P(c^*) : \{b, g\} \to [0, 1]$.

3. *Woman's strategy for each action of the man,* $a_w : \{b, g\} \to \{Accept, Reject\}$.

Notice that, both the social environment and woman holds the same belief about the level of c^*, conditional on which they update their beliefs. Off the path beliefs satisfy the intuitive criterion (Cho and Kreps, 1987).[16]

The following characterizes a partial pooling equilibrium in which a fraction $c^* \in (0, 1)$ of B type men pool with G type men and undertake g. The rest separate and undertake b. G type men always prefer to do g.

Result 1. *(Partial pooling equilibrium with mixed strategy for women) There exists a* $c^* \in (0, 1)$ *where all B type men with* $c \leq c^*$ *undertake g and the rest of the B types undertake b. All G type men undertake g. Social environment and women believe that a fraction* c^* *of the B type men pool. In particular,* $Pr(t_i = G|a_i = g) = \frac{-v(D)}{u-v(D)}$ *and* $Pr(t_i = G|a_i = b) = 0$. *Thus, the beliefs follow Bayes' rule on the equilibrium path. Thus a fraction* $F(c^*)$ *of the B type men undertake g and the rest undertake b. Sequential rationality then implies that women reject if* $a_i = b$ *and accept with a probability q when* $a_i = g$.

[16]The beliefs should be such that for any off equilibrium path information set reached, zero probability should be placed on the types for whom taking the action is equilibrium dominated. Thus, a type will not deviate if the deviation is equilibrium dominated.

Thus, total prevalence of sexual harassment is given by $(1-F(c^*))p$ and total relationships are given by $q[F(c^*)p+(1-p)]$. Notice that by Bayes' rule, $\frac{pF(c^*)}{pF(c^*)+(1-p)} = \frac{u}{u-v(D)}$. This gives that $c^* = F^{-1}(\frac{-(1-p)u}{pv(D)})$. Given the belief, $P(c^*)$, and woman's mixed strategy, q, we can find the cut-off c^* for B type who will be indifferent between doing b and g. The indifference condition is given by $-D = -c^* + q - DP(c^*)$ so that the costs and benefits of doing g for him are equalized. This can be rewritten to give $q = F^{-1}(\frac{-(1-p)u}{pv(D)}) + \frac{v(D)D}{u-v(D)}$. Thus, woman's acceptability of g is dependent on D and p.

Proposition 1. *Male intervention can have two possible effects in the model: It can either lead to an increase in D (social disapproval of those who are perceived to be of B type) or decrease in p (percentage of men who are of B type). Under certain parametric conditions, implications of these two variables on sexual harassment and opposite gender relationships are given below:*

1. *An increase in D increases c^* there by increasing the proportion of B type men who pool with G type men which reduces sexual harassment. However, q decreases which leads to a decline in relationships between men and women due to a decline in women's acceptability of men's offers. Overall sexual harassment decreases, and opposite sex relationships decrease.*

2. *If p decreases, it leads to a decrease in sexual harassment due to a composition effect and also because remaining B type men increase pooling. Women's probability of accepting relationship offers when men approach them with g increases since more men are now good in their class. This leads to an increase in relationships. Overall sexual harassment decreases and opposite sex relationships increase.*

I provide the proofs for the predictions above in appendix C.

The key intuition behind mechanisms above is that a shift in D (social disapproval) or p (proportion B types) can affect B type men's incentive to pool. Increase in D increases B types' incentives to pool but women will take that into account which can reduce q (probability that a woman rejects g), because they suffer from matching with B types. But q may increase, for instance, if D increases so much that the woman's disutility from matching with B types becomes very small (because $v'(D) > 0$). Overall effect on relationships is, thus, ambiguous for increase in D even though sexual harassment decreases. However, I show in appendix C that under certain parametric conditions, a rise in D leads to a fall in relationships. Decrease in p also reduces sexual harassment because the benefit from pooling for the rest of the B type men increases plus the one's who changed their type (from B to G) always do g. Effect on relationships is positive since women take these composition effects into account. Thus, while both the mechanisms predict a reduction in sexual harassment, it is only an increase in D that predicts a decrease in relationships.

Hence testable predictions from the model for male intervention are:

Prediction 1. *An increase in men's perceived social disapproval against sexual harassment (D) reduces sexual harassment and relationships. This occurs because women reduce their acceptance of men's offers. Corollary is that if there is a reduction in relationships, then it means that men's perception of D increased.*

Prediction 2. *An increase in G type men, that is in (1-p), decreases sexual harassment and increases women's relationships. In particular, it increases women's acceptance of men's offers of relationships.*

4 Details of intervention

I collaborated with two colleges to test an awareness intervention with men. I collaborated with one other college to test the female intervention. I collaborated with the NGO Safecity, which specializes in providing sexual harassment awareness trainings and has been active in both urban and rural areas in India since 2013. The sexual harassment awareness training had two main components. One component was informative and had the following main features.

- Legal definition of sexual harassment as per SHWA, 2013, on the role of Internal complaints committees (ICC) in the colleges and legal powers bestowed upon ICCs.
- Principles to detect sexual harassment: this was provided graphically to the students on tablets and by trainers through discussions and skits.
- Situation-based exercises: students were asked whether they thought the situations constituted sexual harassment or not. They were prompted with hints for the correct answer if they answered the questions incorrectly.

The hypothetical situations and principles to detect sexual harassment were developed in consultation with Safecity and other legal experts who have worked in the area of sexual harassment for more than a decade. The informative component of the training helped men understand what legally was sexual harassment. This identifies whether there are awareness constraints that prevent potential perpetrators from understanding what sexual harassment is. Then there were discussions about different types of sexual harassment, for instance in courtship behaviours or even friendships. I theorized that if there are awareness constraints on men then the informative component of the training would alleviate such constraints.

The second component of the intervention was for behavioural change and was intended to be persuasive. I took Safecity's existing sexual harassment awareness training program and tailored it for college students with the help of the trainers. Trainers from Safecity delivered

the training for the intervention. Main features of behavioural component of the training were:

- Detailed in depth discussion about sexual harassment and its impact, including prevalence rates, consent and steps to intervene as a bystander.
- An empathy building section which provided men with anonymous narratives from women explaining how sexual harassment had impacted them in the past, and the prevalence rate of sexual harassment in the same course as them. The idea of empathy building in this case was about perspective taking from the point of view of the harassed.
- Skits and exercises showcasing commonly accepted scenarios of "courtship" that are legally sexual harassment and end up affecting the harassed.

Male intervention was provided in two sessions; in the first sessions there was a 90 minutes workshop and the second session was another doubt session for men only. Each of the sessions was facilitated by a male and a female trainer. These colleges worked under a tight schedule within an academic system controlled by a centralized university and hence the dosage remained low in terms of time. Overall the dosage of the training varied from 3 to 5 hours of intense discussions between trainers and students. However, this also makes the training easily scalable and replicable.

The behavioural change component of the training was to help men understand the impact of the pervasiveness of sexual harassment for victims and why it was an important topic to deal with. For instance, real (anonymous) narratives from female victims of sexual harassment from either the same course in same college or another college were presented to them. One of these narratives is provided in appendix A. A volunteer male student read the narrative and then the trainers led a discussion with the men to understand the effects of seemingly innocuous behaviours. The training is intended to achieve empathy and behaviour change. This also reduced the chances of backlash to the training due to such trainings (Bingham and Scherer (2001)).

The male training between the NGO trainers and the men too place under complete privacy and isolation, giving men a safe space to discuss their thoughts and views openly.[17] Many men continued the discussions well beyond the assigned time period, bringing forward incidents faced by their own female relatives and friends, or sometimes incidents that they were involved in. In addition, the training also included small skits and role plays to prompt discussion. There were discussions about consent, for instance, that 'Only Yes means Yes and everything else is a No' when it comes to consent for sexual activities with another person. The session ended with ways in which men could become a part of the solution rather than

[17]Feedback from trainers after the training revealed that men liked the candid nature of this training.

the problem. Here the trainers usually told men about ways in which they could intervene when they observed sexual harassment. They aimed to help men realise that a discussion on sexual harassment did not limit men as perpetrators but also included them as effective interveners.

The female awareness intervention included only the informative component of the male intervention without any discussion with the NGO trainers. Piloting and discussion with NGOs revealed that women understood the concepts quickly and empathy building was much less relevant for them. Women were given the information to read and could ask the project team in the room any doubts or issues related to it.

5 Experimental Design

The surveys were undertaken by the survey team trained by Abdul Jameel Poverty Action Lab (J-PAL, SA) at the South Asia Center. I collaborated with two colleges for the male intervention and one college for the female intervention.[18]

5.1 Recruitment into the awareness intervention

I, along with the project team, contacted faculty members of the colleges to book a class slot for the survey in advance. Crucially, the class slot was not public knowledge a priori. Hence, the presence or absence of the students was not affected by the content of the survey nor the scheduling. We surveyed the students who were present on the day of the surveys. In this project, I focus only on students who attend college and attend classes. Both men and women, were then told about this being a research collaboration between J-PAL and multiple colleges in the same University. They were also provided with the broad motivation of the project following which they were asked for their informed consent to proceed with the survey.

5.2 Randomization

Unit of randomization for male intervention was a class. A class is a combination of course, year and section.[19] Classes were stratified according to year of study, field of study and sex ratio to provide the male training. Sex ratio is the ratio of baseline enrollment of men to women for each class available from the administrative data. Stratification helps to improve

[18] I was logistically and financially constrained to collaborate with more than one college for the female intervention. Further, this college agreed to collaborate with me much later than the male intervention colleges due to which the timing of surveys is slightly different between male and female intervention colleges.

[19] It also included a medium if the classes were divided by medium. Course means the core subject (like Economics, Maths and so on) which could belong to different fields like science, humanities or a commerce. Sections were usually created for courses with high demand, and a course had a maximum of three sections. Medium refers to the language of instruction which could be Hindi or English. Year was the year of study which could be first, second or third year.

the power of the experiment and control for the class level characteristics that may be correlated with sexual harassment (Glennerster and Takavarasha (2013)). All women in all the classes received the informative component of sexual harassment at this time as well.[20] This was done to remove any under-reporting due to gaps in women's understanding about what constitutes sexual harassment as explained in section 5.5.1.

Sexual harassment awareness intervention for men was at the class level for conceptual reasons. Firstly, sexual harassment awareness training is usually offered to groups rather than one-on-one. In this case, classes were a natural group for delivering the training. The aim of the project was to understand if making men in their environment more aware would reduce women's exposure to sexual harassment. It also helps to increase the power to detect effects on sexual harassment that was collected from women in these classes. I can use women's surveys to elicit sexual harassment perpetrated by men from their own class, which is tightly linked to the training at the class level. It would have been much difficult to ask women about individual men without asking identity of the perpetrator if it was an individual level randomization. Further, offering training to groups facilitated better discussion between the men and the NGO trainers, which pilot testing showed was imperative for a deeper understanding and the reason why it is a favored approach in other settings too.

For providing intervention to women, there were two levels of randomization. Classes were first stratified on year of study, field of study, medium of study and sex ratio at the baseline. Then, 69 classes were divided into high intensity treatment and low intensity treatment. In high intensity treatment classes, 75% of the women were individually randomized to receive information on sexual harassment and in low intensity treatment classes, 25% of the women were individually randomized to receive the female intervention. Figure 2 shows how classes were divided between treatment and control for both the interventions. Class level randomization was done to understand whether the treatment effects on the treated were affected by proportion of their treated peers. Timing of female surveys was delayed by one month and a half.[21]

5.3 Surveys and intervention administration

After the project's introduction, the survey team took the female students (for either male or female intervention) to another private and secluded room. Male students were kept in the original classroom. Both men and women were asked to fill surveys on tablets, separately. The surveys were completely digitized and self-administered by the students via tablets. Crucially, students were placed so that they could not see each others' screens or be influenced by others'

[20] This was a subset of the intervention undertaken for men and provided women with information on how to detect sexual harassment.

[21] This is because the female intervention college came on-board for the project much later than the two colleges for male intervention.

answers, and were monitored by surveyors in the room as depicted in Figure 3 and Figure 4. Students gave informed consent via tablets before filling the survey.[22] Less than 1% students refused the surveys at this point. Pilot surveys were done in a separate college to test whether students understood the questions to answer it themselves. I do not include the data of this pilot college in the paper.[23]

For the male intervention, men in treatment classes were given the information about the project and that the NGO, Safecity wanted to discuss with them various aspects of sexual harassment and also gather their views on it. For the female intervention, women read about the same information on their tablets. For female intervention, randomization of classes was done using STATA. For individual level randomization, women were first given a random card by the survey team that they picked from and were seated on the seat with the same pre-defined number. The women were then handed the tablet after which a randomization code was used to decide whether women in even numbers or odd numbers received the treatment for each class.

5.4 Sample, timeline and balance tests

The two colleges targeted for male intervention had a total of 93 classes. 47 classes out of 93 were randomly assigned to receive the training only for their male students. All classes were re-surveyed for the endline approximately three months after the training. The baseline and training intervention were done towards the beginning of the academic semester (from September until October first week 2019). The endline was done at the beginning of the subsequent semester (January - April 2020).

Since there was a sufficient gap between when the intervention and the endline (relative to the length of academic semester), it was unlikely that Hawthorne effects are a concern. Students were not told that two rounds of the survey were scheduled over one academic year. The faculty members were not aware about it either. Women were asked to recall about men's behaviour in the period between intervention and endline and hence, it was unlikely that men would have changed their behaviour because they were anticipating an endline in the subsequent semester.

A total of 3086 men and women took part in the surveys, 1248 women and 1838 men. In Table 2, I present the balance tests for the combined sample of men and women included in the baseline survey. The sample is balanced on all characteristics. Key features to note

[22] Provided in the appendix

[23] For female surveys, I also provided a helpline number, and Safecity's helpline in case any female respondent needed assistance from professionals after doing the surveys. This is in line with WHO guidelines on surveys on sensitive topics. For male surveys, the same information was provided but they were also given access to the Safecity trainers in case of any further doubts or in case they wanted to talk about their own experiences in treatment classes.

for this sample are that the majority have highly educated parents and are more likely to be from the historically disadvantaged castes (62%). Nearly 25% of them live in a hostel or PG (paying guest accommodation) without any family. About 23% of them have a working mother, which is close to the female labour force participation rate for India. The majority of the students (61%) are from Delhi. The F-stat for joint significance is 1.07 (p value is 0.38), so I can reject the hypothesis that all the variables can jointly explain the assignment to treatment.

Both men and women were surveyed about three months after the intervention, between January to April 2020. The survey team reached 83% of the female baseline population and 80% of the men for a total coverage of 82% of the baseline population for male intervention. In tables E.1 and tables E.2, I show that there was no differential attrition by treatment status or by baseline controls. Women were less likely to drop out than men but it is not differential by treatment status. The p-value for the joint significance test of all baseline variables is 0.89. The survey team could also reach students who were not in the baseline. These are included in the final analysis to help improve power. The balance test for female intervention at the individual woman level is provided in appendix table E.5 and at the class level in table E.6. There are some imbalances at the class level for 3 out of 16 variables which is expected. Nonetheless, I control for all these variables in the regressions. The survey team was able to cover around 86% of the women from baseline with no differential attrition by treatment status of women. A total of 759 women and 1560 men were covered in baseline for female intervention, and we recovered 86% women and 67% men in the endline for female intervention. I show in appendix table E.7 and table E.8 that there was no differential attrition by treatment and that there was no differential attrition by baseline controls except for whether the respondent belonged to the low caste and whether they were originally from Delhi or not.

For generalizability of the results discussed later, I follow (List (2020); Holz et al. (2020)) and report the SANS conditions in appendix D. Next, I discuss how data was collected for various outcomes in the surveys.

5.5 Data

5.5.1 Measuring sexual harassment

Since reporting sexual harassment maybe stigmatized and sensitive, I undertook a number of precautions to collect data on it. These measures were also listed in the ethical protocol that I gained approval for from University of Warwick and IFMR. For this data, I rely on self-reported exposure to sexual harassment from women as perpetrated by men in their own class. A key constraint in studying impact of such a training is connecting the training to

sexual harassment incidence without relying on reports of the men who receive the training. The design helps me overcome this challenge since I asked women about sexual harassment from men in their class. Moreover, women are more likely to recognize a man from their own class, rather than someone from outside which reduces measurement error. These reports of women can then be directly linked to the treatment which was at the class level. This measurement strategy, thus, helps to capture the treatment effects cleanly. Although a legal complaints committee exists in the colleges, students did not reach out to them for complaints and hence, I rely only on survey reports of women.

I discuss below how the measurement strategy overcame several issues that one faces with collecting sexual harassment data.

1. **Selection into the sample**: This was reduced because, as mentioned in section 5.1, students did not know which date or time slot their class was scheduled to be visited by the survey team. Hence, concerns about which women or men fill the survey or not is minimized.

2. **Questionnaire**: I adapt the sexual harassment experiences questionnaire (SEQ henceforth) developed by other researchers (Fitzgerald et al. (1995); Fitzgerald (1988)). This questionnaire has a total of 17 items that are grouped under the categories of gender harassment, unwanted sexual attention, and sexual coercion. Due to logistical constraints, I grouped some of the questions together on the basis of how mild or extreme they were in terms of harm to the harassed. The final questions are provided in the appendix A.1. This reduces fatigue and cognitive load on the women who thus finished the surveys in the limited time allotted to us for the surveys.[24]

3. **Detecting sexual harassment**: To ensure that the male awareness intervention does not create differences in women's awareness of sexual harassment between treatment and control classes, I provided all women in all classes with information on what sexual harassment is. Moreover, chances of differences in detection are minimized in the case of SEQ because of the objective nature of the questions asked.[25]

4. **Under-reporting**: It is still possible, that stigma in treatment classes makes women under-report sexual harassment. (Cullen (2020)) shows that there is no statistically significant difference in reporting of non-partner sexual violence when elicited through a list method and a more direct but tablet-based method which is closer to the method

[24]Further, these questions make answers less subjective or prone to gaps in women's understanding of sexual harassment. In particular, the questions asked women whether a particular incident 'XXX' happened to them rather than whether they were sexually harassed which might be more subjective and prone to information constraints. This necessarily means that I may not be able to cover all kinds of sexual harassment, but I was able to cover maximum number of items in the SEQ.

[25]Indeed as I show in table 5, I do not find any difference in awareness about sexual harassment between women in treatment and control classes

in this paper. I also undertook placebo exercises to show that stigma is not driving the results on sexual harassment reported in section 6.1.1. Further, under-reporting due to the fear of backlash is reduced, since I did not ask women about the identity of the perpetrator in any question.

5. **Privacy of female respondents**: Consistent with what discussions with NGO's revealed, women answer much more truthfully in isolation, atleast from the men. The survey team ensured that all women were taken to a separate isolated room together where they could answer the questions on individual tablets in privacy from college authorities and other students. Figure 4 depicts surveys in progress for women. A team of trained female surveyors was always there in the room so that women could ask them any question they wanted to.[26]

While these are not comprehensive measures to ensure complete truth-telling, these measures are stricter, more comprehensive and complementary to those seen in the literature on sexual harassment (Aguilar et al. (2020), Folke and Rickne (2020), Kondylis et al. (2019)).[27][28] These measures are also more comprehensive than those used by Demographic and Health Surveys to collect data on intimate partner violence for women.

5.5.2 Data for opposite sex relationships

I collected two types of measures to understand effects on relationships with the opposite sex. I first use survey measures to understand the effect on equilibrium outcomes of romantic relationships and friendships. Men and women were surveyed about their romantic partnerships and asked to list friends from their own class. I then create a variable measuring proportion of opposite sex friends they report from their own class and a dummy variable for whether they are dating someone in their own class.[29] The other type of measure is collected through a lab-in-the-field experiment. It is used to understand the effect on men's and women's choices to perform a gender complementary task together. By looking at this separately for men and women, I can study whether effects on the survey measures are due to a change in men's or women's choices.

[26] Female surveyors help in making women participants more comfortable when answering sensitive questions (Aguilar et al. (2020))

[27] Having third party observers in classes to audit sexual harassment reports was not possible, since it would have changed students behaviours.

[28] All women were also told that they had the right to withdraw their data if they wanted to even after submission and they had the first right over the data that they gave to us. I provided all the women with my contact number and that of a resource person at University of Warwick in case they wanted to retract their data. This helped to further increase the students' trust in data privacy. Till now, we have not received any data retraction request from any student

[29] I assume there are no same-sex relationships in this context.

The lab-in-the-field experiment was a between-subjects experiment for which men and women in each class were randomly grouped into mixed or same gender pairs for a class-wide competition.[30] They were then asked to read about a quiz related task that they could either perform with their assigned partner (stick) or alone (switch). The quiz was a combination of 12 questions from female stereotypical or male stereotypical domains. This quiz is an adapted version of tasks used in the literature on gender stereotypes (Bordalo et al. (2019); Coffman et al. (2019); Coffman (2014)). Men and women were then asked simultaneously and privately (on the tablet) whether they wanted to stick or switch with their partner. If they chose to switch, then they solved 6 randomly selected questions from the same quiz. Winners (in teams or individually) were given food vouchers worth 40 INR (40 UK cents). I then compare both men's and women's choice to stick or switch in a mixed-gender versus same-gender group in treatment and control classes. This will help to understand whether the treatment differentially changed incentives for men and women to stick with each other as compared to same gender groups. This is a 2 (male versus female subject) $\times$ 2 (mixed gender versus same gender pair) $\times$ 2 (treatment or control class) design.

A combination of male and female stereotypical questions makes the quiz gender complementary rather than substitutable. This means that the decision to stick or switch away cannot be because either thinks that one sex will be better alone in doing the task. Second, the rewards were such that each member of the pair would receive a food coupon or online voucher if they won. This meant that the decision to stick or switch could not be affected by beliefs that men have a bargaining advantage. In this sense, the experiment will cleanly capture whether treatment affected men and women's choices of switching or sticking with each other on this short-term task. Gender segregation can occur in two ways: women and men do not stick with each other in the experiment or they stick with partners of their own sex more. I use both of these margins as an indicator of a reduced tendency to interact with the opposite gender.[31]

5.5.3 Data for other outcomes

There are a number of challenges in collecting not just data on sexual harassment incidence but also on attitudes towards sexual harassment, behaviour towards the opposite gender, awareness about sexual harassment because of the sensitive nature of the topic. For collecting data on awareness, I developed and piloted questions in the form of hypothetical sexual harassment scenarios which I asked men and women to recognize as sexual harass-

[30] Same gender groups also help with obfuscation (Haaland et al. (2020)), reducing the threat of demand effects.

[31] Further, if the women anticipated any kind of retaliation or backlash from the man they were paired then I can find that women do not switch away from their male partners. This is because if either partner knows what they themselves entered in the survey (to stick or switch), they can infer what their partner entered. In this case then a change in women's preferences can manifest as them sticking with other women rather than switching away from the men.

ment. These scenarios were developed in consultation with NGOs and legal experts working on sexual harassment. I asked respondents three main types of questions to test awareness: a) hypothetical sexual harassment scenarios to identify whether they were sexual harassment or not, b) awareness about legal redressal mechanisms, and c) identifying acceptable courtship behaviours to test whether they chose any sexual harassment behaviour as acceptable or not. Objectivity of these questions helps to alleviate concerns about demand effects. The questions were a mix of sexual harassment or non-sexual harassment scenarios. The detailed questions are in appendix A.2.

Measuring attitudes or beliefs towards sexual harassment in this case is a challenge due to experimenter demand effects (Zizzo (2010)). I asked direct questions and also indirect questions to deal with demand effects. For direct questions, I gave respondents three hypothetical sexual harassment situations. The respondents were told that it was sexual harassment and then, I asked them five sets of questions related to each of the three hypothetical situations. The first set asked them whether they thought that the situation should be legally termed as sexual harassment. This is what I refer to as the direct attitude questions. They are given in appendix section A.4 and were asked to students in only 80% of the classes. For indirect attitude questions, I used two sets of data: google form exercise data and list experiment (Haaland et al. (2020)). In case of list experiment, I masked a sentence on victim blaming attitude within a set of three statements that were contentious but related to sexual harassment and not stigmatized. List experiments help to provide plausible deniability since they ask the participant only about the number of statements that the participants agree with.[32] Within each class, I randomly grouped participants into a list treatment and list control group using their endline survey instrument. The first group read the victim blaming statement masked with three other statements related to sexual harassment. Second group only saw the three statements that were different from victim blaming. Comparing mean number of statements agreed with between the two groups in treatment and control classes gives the differential effect of treatment on victim blaming attitudes. The statements for the list are in appendix A.3

Next, the ICC of the colleges floated a google form during this period inviting students to volunteer or intern for NGOs that work to eradicate sexual harassment and violence against women. These volunteer opportunities were real and the sign-ups were shared with the NGOs as well.[33] I utilized this data to construct a class level variable: share of students who sign up to volunteer for the NGOs. This is available at the class level only since the form was floated through a class Whatsapp group by the complaints committee rather than to individual students.

[32]Most recently, this has been used in (Bursztyn et al. (2020a); Dhar et al. (2018)) to measure stigmatized attitudes in Saudi Arabia and India, respectively.

[33]Since volunteering or internship with NGOs is considered to have considerable returns future labor market prospects, the students had an incentive to sign-up.

Finally, questions about perception of one's peers' support against sexual harassment comprised of both direct questions and indirect measures and are detailed in appendix A.5. Direct questions asked men and women about their perceptions of their classmates' attitudes. I also asked an indirect question on nominations of other students for a class nodal student for taking advise on sexual harassment prevention or reporting. I combined these questions together to form indices to reduce chances of false discovery (Anderson (2008)). Using the same three hypothetical situations that I asked to capture attitudes above, I also asked respondents beliefs about others attitudes towards the law and also about women's propensity to report formally to ICC or informally to classmates and acquaintances for each of the hypothetical situations. These are listed in appendix A.4. I discuss next the empirical strategies I use to examine impact on primary and mechanism outcomes.

All the measures above were collected in the same manner for the female and male interventions.

5.6 Econometric specification

The main econometric specification for understanding the effect of male sexual harassment awareness training on outcomes for both men and women is:

$$Y_{icg} = \beta_1 T_{cg} + \beta_2' X_{icg} + \beta_3' K_{cg} + \alpha_g + \gamma_s + \epsilon_{icg} \tag{1}$$

where i is the student surveyed in the endline survey, c is the class she/he is in, g is the college student i is in, T_{cg} is whether the class c in college g was assigned to receive the male intervention or not, X_{icg} are student characteristics, K_{cg} are class characteristics taken from administrative data, α_g are college level fixed effects, γ_s are strata *(sex ratio×field of study×year of study)* fixed effects following standard practise (Glennerster and Takavarasha (2013)). Y_{icg} is the outcome of interest. Standard errors are clustered at the class level controlling for any correlation in outcomes of students within a class that may be subject to same shocks. Controls are selected by post-double selection LASSO method (Belloni et al. (2014)). If the baseline controls are missing for some individuals, then I control for a dummy variable indicating whether the variable was missing for the respondent or not. β_1 captures the intent to treat effect of the training on student i's outcome.

The empirical specification for the lab-in-the-field experiment takes the following form:

$$Y_{mcg} = \beta_1 T_{cg} + \beta_2 Mixed_Gender_{mcg} + \beta_3 T_{cg} \times Mixed_Gender_{mcg} + \beta_4' X_{mcg} \tag{2}$$

$$+\beta_5' K_{cg} + \alpha_g + \gamma_s + \epsilon_{mcg}$$

In this equation, I look at the binary decision of male student m, in class c in college g to stick to doing the task with his randomly assigned partner ($Y_{mcg} = 1$) rather than opting to

do it alone ($Y_{mcg} = 0$). $MixedGender_{mcg}$ is a dummy equal to 1 if m was assigned to a mixed gender group. The omitted category is the same gender group in control classes. X_{mcg} are student level controls and K_{cg} are class level controls. Standard errors were clustered at the class level. I run a similar regression for women.

I will show results from the female intervention using the following specification that exploits both levels of randomization for female intervention.

$$Y_{ic} = \beta_1 Female_treatment_{ic} + \beta_2 High_Intensity_c + \tag{3}$$

$$\beta_3 High_Intensity_c \times Female_treatment_{ic} + \gamma_s + \beta_4 X_{ic} + \epsilon_{ic}$$

Y_{ic} is the relevant dependent variable, $Female_treatment_ic$ is a dummy that takes value 1 if the woman i was assigned to the treatment, and 0 if she was not and $High_Intensity_c$ is a dummy that takes value 1 if class c was assigned to the high intensity treatment. β_3 is the difference in outcome between someone who is treated in the high intensity class versus someone who was not treated in the same class, β_2 is the effect of being an untreated woman in a high intensity class as against someone who is untreated in a low intensity class and β_1 gives the effect of being treated in a low intensity class as compared to someone who is untreated in a low intensity class. γ_s are strata fixed effects. Standard errors are clustered at the class level and controls are selected by post-double selection LASSO method.

6 Results

In what follows, I will first discuss the results of the male intervention for primary and mechanism outcomes and then discuss the female intervention for comparison.

6.1 Results for the male intervention

6.1.1 Impact on sexual harassment

I first show results on the effects of the training on sexual harassment reported by women. I utilize the question on sexual harassment that I asked women about as perpetrated from men in their own class. As explained before, this is more tightly linked to the treatment and helps to overcome issues related to identity of the harasser. I asked women about different types of sexual harassment; mild, intermediate and extreme as mentioned in section 2. I look at the effects of the training on these three different types of events and then look at an overall index -same class index- that combines all types together. The results are shown in table 3. Training reduces sexual harassment perpetrated by men from treatment classes by 0.06 sd as reported by their female classmates. I also look at the effects on different types of sexual harassment. Training reduces incidence of extreme forms of sexual harassment perpetrated

by men from training classes by 1.05 p.p (or 0.125 sd) at 1% level of significance. Notice 1% of women in control group report being harassed physically by men in their class over a period of three months preceding the survey. Thus, the training was highly effective in eliminating arguably the more harmful forms of sexual harassment from the treated men.

The results are robust to randomization inference, thereby allaying concerns about the low incidence of extreme forms of harassment in control group. The results are also robust to multiple hypothesis testing. In a placebo exercise, I show in appendix table E.10 and table E.11, that there are no such negative effects on women's reporting of sexual harassment from men in a different class or men from outside college. I find a marginally significant increase in women's reporting of extreme forms of harassment form men outside the college. This is because women's relationships with men outside their college increased after the treatment. Thus, this gives greater confidence that the effect is due to the treatment affecting men's behaviour rather than women's reporting behaviour.[34] The results are also robust to alternative samples over which I created the index. This is shown in appendix table E.9.

Using the estimates for extreme forms of sexual harassment which went down by 1.1 p.p over a control mean of 0.01, translates into 51 fewer women who face extreme forms of sexual harassment over an entire academic year (which includes two semesters). This should be taken as a lower bound on actual number of incidents of sexual harassment since the outcome captures only whether the woman faced sexual harassment or not. Overall, an effect of 0.06 sd after a training of 3 to 5 hours is a strong effect and comparable with effects of community based training programs on intimate partner violence. I provide comparison of the effect sizes in other experimental studies in figure 5 and figure 6. With the caveat that these studies focus only on IPV and other crimes, my results for overall sexual harassment are close to Green et al. (2020) who show negative effects of a movie screening program in Uganda on domestic violence reported by women.[35] For extreme forms, my results which show 100% decline is the highest magnitude as compared to other papers.

This training is easy to incorporate into a standard curriculum and hence scalable as well. The results show that a training that combines both an informative (awareness component) and a persuasive component (empathy building) can help deter sexual harassment in a high prevalence context. In particular, the training reduces quite effectively the most extreme forms of sexual harassment that potentially impose highest costs on the society and for the harassed.[36]

[34] I show in the appendix table E.22 that if I were to include sexual harassment from men in any environment, that is within class, outside class or outside college, then there are no effects on overall sexual harassment but that is because of the increase in sexual harassment from outside the college.

[35] I first converted there estimates to reflect effects over 3 month period.

[36] The economic costs of sexual harassment in the workplace, Deloitte Report (2019).

6.1.2 Opposite sex relationships

Next, I study effects of the training on opposite sex relationships in treatment versus control classes. If the perceived social disapproval (D) becomes higher, then it is possible that relationships decline because women reduce their acceptability of such relationships (theoretical prediction 1). On the other hand, if only the proportion of G types increases, then there can be an increase in relationships (theoretical prediction 2). Results on relationships will help to distinguish between the two mechanisms.

In table 4, I report regression results of effects on equilibrium outcomes in columns 2 and 3 and for choices in the lab-in-the-field experiment in columns 4 and 5. The upper panel provides the effects on men and the lower panel on women. I find that the training reduces opposite sex romantic partnerships by 1.3 p.p in treatment classes in column 2. This corresponds to a 64% reduction on average as compared to control mean of 2%. The coefficient is reassuringly similar for men although the effect for men is insignificant. There is a negative but insignificant effect on friendships in column 3. As mentioned before columns 2 and 3 are equilibrium outcomes. This is consistent with theoretical prediction 1, that is, an increase in D. Ex-ante, it is unclear whether these effects originate from women or men and which one will be more dominant. The theoretical framework shows that the effects on relationships are due to a change in women's choices. I use results from lab-in-the-field to study this next.

Since the task that each pair in the lab-in-the-field experiment had to solve was gender complementary, there should, thus, be atleast as high a tendency to stick to one's partner in a mixed gender pair as that in a same gender pair in control classes. The reasons are listed in section 5.5.2. This is what I find in Figure 7 for control group proving that respondents did take stereotypical nature of the task into account. It also shows that the treatment increased men's tendency to stick with women in mixed gender pairs as compared to the control group with no such effect for women (although this is statistically insignificant). However, women stick with each other much more than with men in treatment versus control classes (74% versus 63%) indicating breakdown of relationships on account of women's choices.

In column 4 of table 4, I find that women's preference for cooperation within same gender pairs increases due to the training by approximately 14 p.p. (37% increase over control). However, I cannot detect any effect on men's choices. Combining the survey measures and lab experiment, I create an opposite sex relationship index using (Anderson (2008)) in column 1.[37] I find that there is an overall decrease in opposite sex relationships index by approximately 0.13 sd for women and an overall insignificant and small effect on index reported by men. Thus, the sexual harassment awareness training has negative effects on such relationships

[37]Since the lab-in-the-field experiment was a between subjects experiment, I had to impute the missing values for those who were not assigned to a particular group. I used the KLK method to impute these values (Kling et al. (2007))

suggestively on account of a change in women's choices. The results are also robust to multiple hypothesis testing and randomization inference for the index.

These results suggest that since it was women who most likely changed their choices away from men atleast in the lab-in-the-field experiment, then the results on relationships are more likely due to women's lack of acceptance rather than men's lack of offer of such relationships. To further show that men's offers did not change, I also asked women in a survey question on whether they were approached by anyone in their class to form a romantic relationship with them, and I found no treatment effect on this measure in table E.18. I further show that the treatment effects on opposite sex relationships are stronger for respondents in their first year of college who would have less information about each other, in particular about the type of men in table E.13. Lastly, I show in table E.19 that women's relationships with men outside the class increased showing that it was a change in the treated men's behaviour that was affecting women's behaviour. In particular, women are not reducing relationships with all men. From the lens of the model, the results for sexual harassment and opposite sex relationships are consistent with an increase in men's perception of social disapproval (D) rather than men's intrinsic attitudes. Lack of information about men's types plus the higher pooling of B types reduces women's preference to form relationships with men. Thus, I find support for theoretical prediction 1.

6.1.3 Mechanism outcomes

In this section, I study the main mechanisms highlighted in section 1 to explain the effects I find on sexual harassment and opposite sex relationships through the lens of the theoretical framework.

In table 5, I present results for the effect of the training on men's awareness about sexual harassment. The results help shed light on whether there is any lack of awareness about sexual harassment for men in recognizing what behaviours constitute sexual harassment and whether the training helps to alleviate it. I find that men in treatment classes are 0.09 sd (in column 1) more aware approximately three months after the training than men in control classes. Awareness about legal mechanisms increases by 107% (column 6), awareness about ambiguous sexual harassment situation increases by 12% (column 4). It is reassuring that the treatment is able to affect men's awareness nearly three months after the training after a dosage of three to five hours. There is no overall effect on women's awareness about sexual harassment in treatment classes as compared to women in control classes though.[38] This is likely because all women were provided with information on sexual harassment at the baseline, to which the male intervention could not add much more. Within the theoretical framework,

[38]Although women too become more aware about ICC in training classes by 6.3 p.p (column 6) of a base of 20%. The spillover effect on women within treatment classes shows that there may have been increased interaction between men and women about ICC but it is not strong enough.

an increase in awareness can have two effects: It will induce some unaware G types to change behaviour (directly reducing sexual harassment), and/or induce some B types to change their attitudes (decrease in p).

This begs the question if increased awareness was not the only driver of results. Firstly, as mentioned above, the hypothesized effects of awareness would lead to an increase in relationships within the theoretical framework. On the other hand, in my results, I find the opposite. Secondly, I show in figure 8, that awareness increased significantly for intermediate and mild forms of sexual harassment rather than extreme. But since sexual harassment of extreme events went down significantly instead shows that the mechanism cannot be awareness atleast for extreme events.

Next, I test whether the training changed men's beliefs about how costly it is to perpetrate sexual harassment. One argument given for effectiveness of sexual harassment awareness training is deterrence (SHWA, 2013). In particular, ICC's are the main formal and mandatory bodies set up to prevent and redress sexual harassment in colleges and workplaces. However, the role of bystanders and informal institutions is also documented to have a deterrent effect on crime (Nagin et al. (2013)). Withing colleges, these informal institutions are sanctions by classmates or friends who can intervene, call out sexual harassment or provide informal support if someone is a target of sexual harassment. I consider the latter type of deterrence as social disapproval of B type men by peers. Since the training was provided to all men in treatment classes, there could potentially be greater change in social disapproval against sexual harassment than legal costs imposed by the ICC.[39] I collected both men's and women's beliefs or perceptions about both kinds of mechanisms.

I asked multiple questions on respondents perception of social and legal costs to sexual harassment. I combine these questions in indices separately for both types of costs. I find that there is a very weak effect on perception of legal costs to sexual harassment in column 2 in table 6 for men. However, there was a significant decline in women's perception of legal costs to sexual harassment by approximately 0.08 sd (column 1, panel B). In a stark contrast to this, I find that perception of social costs to sexual harassment increased strongly for men by 0.05 sd (significant at 5% level) in column 1. I also provide results for the components of the indices in detail in appendix tables E.15 and table E.16. Additionally, I asked men in 80% of the classes to report their perceived probability that a woman in their class will report about three hypothetical sexual harassment scenarios to their own classmates or acquaintances and to the ICC and their beliefs about other men and women's acceptability of the three situations as sexual harassment legally. The results are in appendix table E.17, table E.20, and table E.21. I find that for each of the situations, men report higher

[39] As explained earlier, ICCs or internal complaints committees were new to these colleges and there was a significant amount of learning-by-doing reported by members in these committees at the time of the pilot. This meant that not only men and women were unaware about these ICCs in their colleges but also would not trust their effectiveness.

probability that women will report to their classmates rather than to ICC. Their perception about other men's and women's acceptance of the law increased significantly while it did not change for women. The combination of these results is important. In particular, it shows that deterrence against sexual harassment likely came from increases in social disapproval rather than from perceived legal costs. From the lens of the theoretical framework, an increase in men's perception of social image costs to sexual harassment (D) increases which is consistent with results on sexual harassment and relationships.

Next, I test the effect of the training on men's own direct attitudes towards sexual harassment. I first show the effects on attitudes collected using direct questions from men in table 6. Results for men indicate that it is possible that intrinsic attitudes of men increased however since these questions were asked directly, they can suffer from demand effects. This would also be particularly severe for attitudes on sensitive topics like sexual harassment. As explained in section 5.5, I used a list experiment and a google form exercise to measure changes in intrinsic attitudes or beliefs. Both the list experiment and sign-up exercise help to understand if beliefs or attitudes towards sexual harassment changed. I assume here that if men do not empathize with the issue of sexual harassment or blame the victims for sexual harassment then it indicates that their intrinsic attitude towards sexual harassment perpetration did not change either. I show the results for both of these variables in table 6. I find that there was no effect on either men's victim blaming attitudes (0.001) or the google form sign ups (-0.002). The coefficients are also very small relative to the mean. MDE for list experiment was 0.03 sd and for google form exercise was 0.2 sd which is at most what has been found in other literature for attitudes. This is in contrast to (Dhar et al. (2018)) that find that a gender sensitization program improved students' gender attitudes in India by 0.179 sd. This could be because sexual harassment attitudes may be particular difficult to change and that I focus on older students for whom such attitudes maybe less malleable.

Overall these results indicate that the training increases men's perceptions of social disapproval of sexual harassment without there being a detectable change in their intrinsic attitudes. This is similar to the insight in (Bursztyn and Yang (2021); Bursztyn et al. (2020a)) that individual perceptions might not correctly reflect beliefs or actual attitudes of those around but can affect own actions nonetheless.

Next, I discuss results from awareness intervention for women to help compare results with those of male intervention.

6.2 Comparison with female sexual harassment awareness intervention

Interpreting female intervention results needs caution. This is because providing women with information on sexual harassment can change their reporting due to increased awareness or knowledge, however, it can also induce changes in women's behaviour which affects their actual

exposure to sexual harassment. It is thus difficult to disentangle these effects. Comparing treated women in high intensity and low intensity classes helps to overcome this challenge to some extent. Hence I report F test of equality of coefficients for these women to understand the effect of the class level treatment.

6.2.1 Impact on sexual harassment

I show results from estimating the regression in equation 3 in table 7. First, the estimate for β_3 is negative for all types of sexual harassment although insignificant. When I combine the responses into a sexual harassment from same class index, I continue to find that this coefficient is negative (0.17 sd). I cannot reject a null effect of the treatment for women assigned to individual level treatment in a high intensity treatment class. I find that the effect on untreated women in high intensity class is positive except for extreme events. Finally, the coefficients are close to zero for the effect of the individual treatment on women in a low intensity class. In all the cases, I cannot reject a null effect. I also test whether $\beta_2+\beta_3-\beta_1 = 0$, but do not detect any statistically significant differences between the two. I can rule out an effect of the size of 0.1 or above with the class level female intervention. It is still inconclusive to say whether the male intervention is more effective since it shows a lower effect.

6.2.2 Opposite sex relationships

Next, I show results from a similar specification as above for opposite sex relationships index in table 8. Firstly, treated women in high intensity classes are for most part less likely to prefer interacting, cooperating or forming a relationship with men in their class. However coefficient on the interaction term is not significant for any of the outcomes. Individual level female treatment alone reduces women's romantic relationships with men in their class by nearly 4 p.p. This is a very strong effect as compared to the male treatment. Compared to the control group this is a complete reduction in romantic relationships but I do not find any such detectable effects for women in high intensity class. An F-test shows that indeed treated women in low intensity classes have lower romantic relationships with men than treated women in high intensity classes. For friendships, there is a marginal negative effect on friendships of 6 p.p for untreated women in high intensity classes. However the F-test rules out that there is any detectable difference in the results for treated women in high and low intensity classes. I can rule out effect size of 0.1 or more.

6.2.3 Mechanism outcomes

Finally, I look at mechanisms for both men and women together in a coefficient plot in figure 9. For men in particular, I find there is a slightly positive effect on perceived social costs of sexual harassment (significant at 10% level) but it is not strong enough. Female treatment has a

precisely estimated null effect on awareness of men which is also intuitive since the treatment was given to the women. An important difference, thus, here is that even if we believe that perception of peer support increased against sexual harassment even due to the class level treatment of women, it may have limited effects on sexual harassment if men do not know what 'non sexually harassing' behaviours are.

6.3 Long run Results

6.3.1 Main outcomes

Here I discuss results from the long run surveys undertaken with the same sample for male intervention. I show the results for sexual harassment and romantic relationships in table 9 using only the long run sample. I find that the male treatment leads to a marginally significant decline in mostly intermediate forms of harassment like online stalking, staring, stalking but that the overall effects on sexual harassment are insignificant. Long run data is noisier especially because women were asked to recall about sexual harassment for before the pandemic. This might lead to noisier estimates even though the coefficient on overall index is higher than before. Romantic relationships between men and women in the same class continue to be lower while women still report being in a relationship with men outside the class over the long run even after they have graduated from college. In table 10, I show the results for all the rounds of survey together to increase power and also to get an overall picture. I find that overall the effects on extreme forms of harassment continue to hold strong. Romantic relationships in treatment classes continue to be approx 2 p.p lower than those in control classes. These results indicate the persistent impact of the treatment on opposite sex relationships while the results on sexual harassment are stronger in the short to medium run and noisier for long run. This could be an artifact of the reliance of recall data for sexual harassment from women.

I also show results for the mechanism outcomes for women and they are in line with the short run results. In particular, women's awareness, perceived informal or formal support towards sexual harassment did not change in the long run. Interestingly, women's attitudes didnot change in the long run which I tested for using a donation experiment and using a similar list experiment as in the short run to measure victim blaming attitudes of women. These results are in appendix table E.23.

7 Discussion and conclusion

In this paper, I provide experimental evidence on the impact of sexual harassment awareness training for college students on sexual harassment outcomes for women and on relationships between men and women. I highlight in particular the mechanisms behind effects of this

training. The results of this paper directly inform policy makers and law makers around the world who have advocated for sexual harassment awareness training. The paper shows that the training for men helps to reduce sexual harassment for women. Most of such training is provided in groups, typically with one's peers. I show that this is key to its effectiveness in reducing sexual harassment. In particular, such training can affect men's perception of their peers' attitudes (perceived social disapproval) thereby inducing them to change behaviour. But in a framework where women used men's behaviour to screen between good and bad type men, women are can become also constrained in screening between these two types after the training. This is key, because I do not detect effects on men's intrinsic attitudes towards sexual harassment using measures that alleviate experimenter demand effects which may be severe in any self-reported data on attitudes.

In a collaboration with an NGO, Safecity, I provided sexual harassment awareness training to men in randomly selected classes at collaborating colleges in Delhi. This training informed men about laws against sexual harassment, the definition of sexual harassment and legal procedures in their college for handling sexual harassment complaints. It also helped men to think about how sexual harassment affects women, that is, to empathize with them. I find that the training successfully increased men's awareness about sexual harassment and legal procedures for dealing with it by 0.09 sd which indicates large awareness constraints about sexual harassment. The training successfully reduces overall sexual harassment by 0.06 s.d.. I used a lab-in-the-field experiment and survey measures to study effects of the training on other types of relationships between men and women. Men and women are less likely to form romantic relationships with each other largely due to a shift in women's preferences. Overall, there is more gender segregation due to a change in women's preferences.

To the extent that the inter-personal environment of educational institutions mimics that of workplaces, this has important repercussions for workplace relationships. A more thorough understanding of how this training might affect group productivity, cooperation and performance for workplaces is needed. I cannot reject null effects of the female intervention on sexual harassment or on opposite-sex relationships. Studying female intervention is complicated since it can affect both awareness (without change in actual incidence) and actual incidence of sexual harassment.

Sexual harassment awareness training for men took a total of 3 to 5 hours in my setting and hence, was a relatively short intervention in comparison to other studies that use attitude change interventions. The reduction in sexual harassment gives confidence about scalability of the intervention. Journalists have claimed that sexual harassment and abuse on campus can cost universities millions of dollars not only through lawsuits, but also through reduced

alumni donations and future admissions.[40] The inter-personal setting of universities makes this study relevant for workplaces too.[41]

Recent work on gender inequality shows that attitude change programs targeted at men may be important, since they have the power to act on the knowledge (Dhar et al. (2018)) they acquire from such programs. Sexual harassment is an area where involvement of men may be crucial not only because men are the majority of the perpetrators but also because they may have more power to induce behaviour change in others too. A key constraint on men then might be how disapproving against undesirable behaviours they perceive their social environment to be.

[40] *Campus Sexual Assault can cost Universities millions*, Forbes, January 2015.

[41] Workplace safety for women is a crucial factor in improving women's labour market participation, engagement and aspirations (Jayachandran (2020); Azmat et al. (2020); Chaudhary et al. (2014); Sudarshan and Bhattacharya (2009)).

References

Abramsky, T., K. Devries, L. Kiss, J. Nakuti, N. Kyegombe, E. Starmann, B. Cundill, L. Francisco, D. Kaye, T. Musuya, et al. (2014): "Findings from the SASA! Study: A cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda," *BMC medicine*, 12, 122.

Aguilar, A. A., E. Gutierrez, and P. S. Villagran (2020): "Benefits and Unintended Consequences of Gender Segregation in Public Transportation: Evidence from Mexico City's Subway System," .

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

Anderberg, D., N. Mantovan, and R. M. Sauer (2018): "The Dynamics of Domestic Violence: Learning about the Match," .

Anderberg, D. and H. Rainer (2013): "Economic Abuse: A Theory of Intrahousehold Sabotage," *Journal of Public Economics*, 97, 282–295.

Anderson, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103, 1481–1495.

Antecol, H. and D. Cobb-Clark (2003): "Does Sexual Harassment Training change Attitudes? A View from the Federal Level," *Social Science Quarterly*, 84, 826–842.

——— (2006): "The Sexual Harassment of Female Active-duty Personnel: Effects on Job Satisfaction and Intentions to Remain in the Military," *Journal of Economic Behavior & Organization*, 61, 55–80.

Ashraf, N., N. Bau, C. Low, and K. McGinn (2020a): "Negotiating a Better Future: How Interpersonal Skills Facilitate Interg-enerational Investment," *The Quarterly Journal of Economics*, 135, 1095–1151.

Ashraf, N., F. Erica, A. Voena, and R. Ziparo (2020b): "Maternal Mortality Risks and Spousal Differences in Desired Fertility," *Working Paper*.

Azmat, G., V. Cuñat, and E. Henry (2020): "Gender Promotion Gaps: Career Aspirations and Workplace Discrimination," *CEPR Discussion Paper No. DP14311*.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2020): "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa," *American Economic Journal: Applied Economics*, 12, 210–59.

Banerjee, A., E. L. Ferrara, and V. Orozco (2019): "Entertainment, Education, and Attitudes toward Domestic Violence," *AEA Papers and Proceedings*, 109, 133–37.

Banerjee, A. V., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2018): "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets," *Available at SSRN 3245656*.

Basu, K. (2003): "The Economics and Law of Sexual Harassment in the Workplace," *Journal of Economic Perspectives*, 17, 141–157.

Beaman, L. and A. Dillon (2018): "Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali," *Journal of Development Economics*, 133, 147–161.

Beaman, L., N. Keleher, and J. Magruder (2018): "Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi," *Journal of Labor Economics*, 36, 121–157.

Beaman, L. and J. Magruder (2012): "Who Gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 102, 3574–93.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *The Review of Economic Studies*, 81, 608–650.

Bénabou, R. and J. Tirole (2006): "Incentives and Prosocial Behavior," *American economic review*, 96, 1652–1678.

Bernheim, B. D. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841–877.

Bingham, S. G. and L. L. Scherer (2001): "The Unexpected Effects of a Sexual Harassment Educational Program," *The Journal of Applied Behavioral Science*, 37, 125–153.

Bloch, F. and V. Rao (2002): "Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India," *American Economic Review*, 92, 1029–1043.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019): "Beliefs about gender," *American Economic Review*, 109, 739–73.

Borker, G. (2017): "Safety First: Perceived Risk of Street Harassment and Educational Choices of Women," *Job Market Paper, Department of Economics, Brown University*.

Buchmann, N., E. M. Field, R. Glennerster, S. Nazneen, and X. Y. Wang (2021): "A Signal to End Child Marriage: Theory and Experimental Evidence from Bangladesh," .

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2018): "Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia," Tech. rep., National Bureau of Economic Research.

——— (2020a): "Misperceived social norms: Women working outside the home in saudi arabia," *American Economic Review*, 110, 2997–3029.

Bursztyn, L., I. K. Haaland, A. Rao, and C. P. Roth (2020b): "Disguising Prejudice: Popular Rationales as Excuses for Intolerant Expression," .

Bursztyn, L. and R. Jensen (2015): "How does peer pressure affect educational investments?" *The Quarterly Journal of Economics*, 130, 1329–1367.

Bursztyn, L. and D. Y. Yang (2021): "Misperceptions about Others," .

Calvi, R. and A. Keskar (2020): "'Til Dowry Do Us Part: Bargaining and Violence in Indian Families," *Working Paper*.

Canen, N., M. O. Jackson, and F. Trebbi (2019): "Endogenous Networks and Legislative Activity," *Available at SSRN 2823338*.

Chaudhary, R., S. Verick, et al. (2014): *Female labour force participation in India and beyond*, ILO New Delhi.

Cho, I.-K. and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

Coffman, K. B. (2014): "Evidence on self-stereotyping and the contribution of ideas," *The Quarterly Journal of Economics*, 129, 1625–1660.

Coffman, K. B., M. Collis, and L. Kulkarni (2019): "Stereotypes and Belief Updating," *HBS Working Paper No. 19-068*.

Comola, M. and S. Prina (2014): "Do Interventions Change the Network? A Dynamic Peer Effect Model Accounting for Network Changes," *Unpublished manuscript, Paris School of Economics*.

Cullen, C. A. (2020): "Method Matters: Underreporting of Intimate Partner Violence in Nigeria and Rwanda," *World Bank Policy Research Working Paper*.

Dahl, G., A. Kotsadam, and D.-O. Rooth (2020): "Does Integration Change Gender Attitudes? The Effect of Randomly Assigning Women to Traditionally Male Teams," *The Quarterly Journal of Economics, Forthcoming*.

Dahl, G. B. and M. M. Knepper (2021): "Why is Workplace Sexual Harassment Underreported? The Value of Outside Options Amid the Threat of Retaliation," .

Dean, J. T. and S. Jayachandran (2019): "Changing Family Attitudes to Promote Female Employment," in *AEA Papers and Proceedings*, vol. 109, 138–42.

DellaVigna, S., J. A. List, and U. Malmendier (2012): "Testing for Altruism and Social Pressure in Charitable Giving," *The quarterly journal of economics*, 127, 1–56.

DellaVigna, S., J. A. List, U. Malmendier, and G. Rao (2016): "Voting to Tell Others," *The Review of Economic Studies*, 84, 143–181.

Dhar, D., T. Jain, and S. Jayachandran (2018): "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," *NBER Working Paper, 25331.*

Duflo, E. (2012): "Women Empowerment and Economic Development," *Journal of Economic literature*, 50, 1051–79.

Dupas, P., A. Keats, and J. Robinson (2019): "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya," *The Economic Journal*, 129, 273–310.

Edmonds, E. V., B. Feigenberg, and J. Leight (2020): "Advancing the Agency of Adolescent Girls," .

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

Field, E., S. Jayachandran, and R. Pande (2010): "Do Traditional Institutions constrain Female Entrepreneurship? A Field Experiment on Business Training in India," *American Economic Review*, 100, 125–29.

Field, E., S. Jayachandran, R. Pande, and N. Rigol (2016): "Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship?" *American Economic Journal: Economic Policy*, 8, 125–53.

Fitzgerald, L. (1988): "The Dimensions and Extent of Sexual Harassment in Higher Education and the Workplace," *J Vocat Behav*, 32, 152–175.

Fitzgerald, L. F., M. J. Gelfand, and F. Drasgow (1995): "Measuring Sexual Harassment: Theoretical and Psychometric Advances," *Basic and applied social psychology*, 17, 425–445.

Fitzgerald, L. F. and S. L. Shullman (1993): "Sexual harassment: A Research Analysis and Agenda for the 1990s," *Journal of Vocational Behavior*, 42, 5–27.

Folke, O. and J. K. Rickne (2020): "Sexual Harassment and Gender Inequality in the Labor Market," *CEPR Discussion Paper No. DP14737*.

Glennerster, R. and K. Takavarasha (2013): *Running randomized evaluations: A practical guide*, Princeton University Press.

Green, D. P., A. M. Wilke, and J. Cooper (2020): "Countering Violence against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda," *Comparative Political Studies*, 0010414020912275.

Greif, A. (1989): "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders," *The journal of economic history*, 49, 857–882.

Gulesci, S., M. Puente Beccar, and D. Ubfal (2021): "Can Youth Empowerment Programs Reduce Violence against Girls during the COVID-19 Pandemic?" *Journal of Development Economics*, 153.

Haaland, I., C. Roth, and J. Wohlfart (2020): "Designing Information Provision Experiments," *CEBI Working Paper Series, Working Paper 20/20*.

Haushofer, J., C. Ringdal, J. P. Shapiro, and X. Y. Wang (2019): "Income Changes and Intimate Partner Violence: Evidence from Unconditional Cash Transfers in Kenya," *National Bureau of Economic Research Working Paper No. 25627*.

Hersch, J. (2011): "Compensating Differentials for Sexual Harassment," *American Economic Review*, 101, 630–34.

——— (2018): "Valuing the Risk of Workplace Sexual Harassment," *Journal of Risk and Uncertainty*, 57, 111–131.

Holz, J. E., J. A. List, A. Zentner, M. Cardoza, and J. Zentner (2020): "The $100 Million Nudge: Increasing Tax Compliance of Businesses and the Self-Employed using a Natural Field Experiment," Tech. rep., National Bureau of Economic Research.

Jäckering, L., T. Gödecke, and M. Wollni (2019): "Do Agricultural Interventions Influence Network Formation? Insights from a Randomized Experiment in Kenya," *Working Paper*.

Jayachandran, S. (2015): "The Roots of Gender Inequality in Developing Countries," *Annual Reviews of Economics*, 7, 63–88.

——— (2020): "Social Norms as a Barrier to Women's Employment in Developing Countries," .

Karing, A. (2018): "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone," *University of California, Berkeley*, 2.

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

Kondylis, F., A. Legovini, Arianna amd Zwager, and K. Vyborny (2019): "Sexual Harassment on Transit: Safe Spaces and Stigma," *World Bank Working Paper, 9269*.

Kotsadam, A. and E. Villanger (2020): "Jobs and Intimate Partner Violence-Evidence from a Field Experiment in Ethiopia," .

Lindo, J. M., P. Siminski, and I. D. Swensen (2018): "College Party Culture and Sexual Assault," *American Economic Journal: Applied Economics*, 10, 236–65.

List, J. A. (2020): "Non est disputandum de generalizability? a glimpse into the external validity trial," *NBER Working Paper*.

Macchiavello, R. and A. Morjaria (2015): "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review*, 105, 2911–45.

Mayer, A. and S. L. Puller (2008): "The Old Boy (and Girl) Network: Social Network Formation on University Campuses," *Journal of Public Economics*, 92, 329–347.

McKelway, M. (2020): "Women's Employment in India: Intra-household and Intra-personal Constraints," .

Miller, C., J. Peck, and M. Seflek (2019): "Integration Costs and Missing Women in Firms," Tech. rep., National Bureau of Economic Research.

Nagin, D. S. et al. (2013): "Deterrence: A Review of the Evidence by a Criminologist for Economists," *Annual Review of Economics*, 5, 83–105.

Roehling, M. V. and J. Huang (2018): "Sexual Harassment Training Effectiveness: An Interdisciplinary Review and Call for Research," *Journal of Organizational Behavior*, 39, 134–150.

Ru, H. and R. M. Townsend (2020): "Enhanced Informal Networks: Costly State Verification and the Village Fund Intervention," *Available at SSRN*.

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 116, 681–704.

Sudarshan, R. M. and S. Bhattacharya (2009): "Through the Magnifying Glass: Women's Work and Labour force participation in urban Delhi," *Economic and Political Weekly*, 59–66.

Talboys, S. L., M. Kaur, J. VanDerslice, L. H. Gren, H. Bhattacharya, and S. C. Alder (2017): "What is Eve Teasing? A Mixed Methods Study of Sexual Harassment of Young Women in the Rural Indian Context," *Sage open*, 7.

Zhu, M. (2019): "Job Networks through College Classmates: Effects of Referrals for Men and Women," *Working Paper, Syracuse University*, 1.

Zizzo, D. J. (2010): "Experimenter Demand Effects in Economic Experiments," *Experimental Economics*, 13, 75–98.

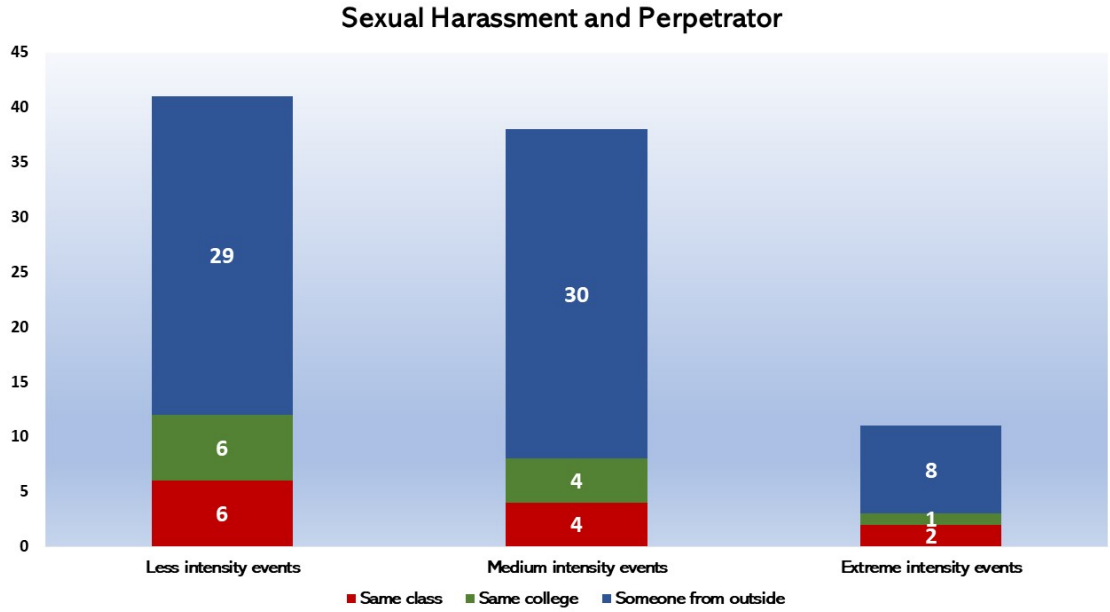


Figure 1: Perpetrators of sexual harassment as reported by women

The figure above shows the percentage of women who faced sexual harassment from men in different categories. The men are categorized into three groups, a) someone from outside college, b) someone in college but not same class as the female respondent and c) someone from same class as the female respondent.

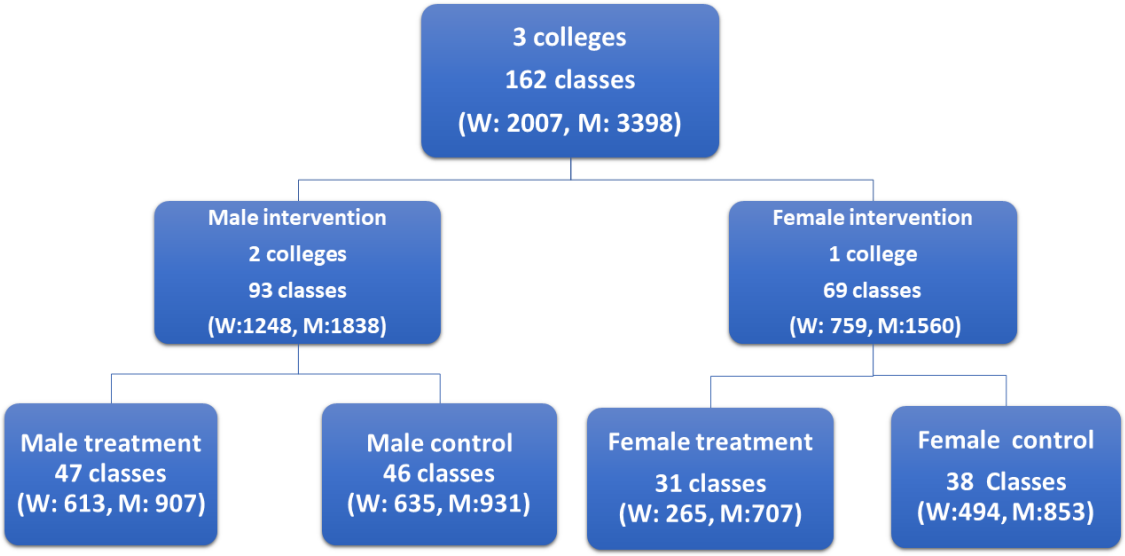


Figure 2: Overall Design



Figure 3: Male training

Men were taken to rooms with projectors for the training. In this picture two trainers from Safecity deliver the training session to men in a treatment class. All men from the same class received the training together.



Figure 4: Female surveys

Women were taken to a separate room to fill the 'female surveys' on tablets. They were seated at a distance from each other to ensure privacy. One female member from the survey team always remained in the room to answer questions and resolve technical difficulties during the survey. This picture was taken with the consent of female students in the picture.

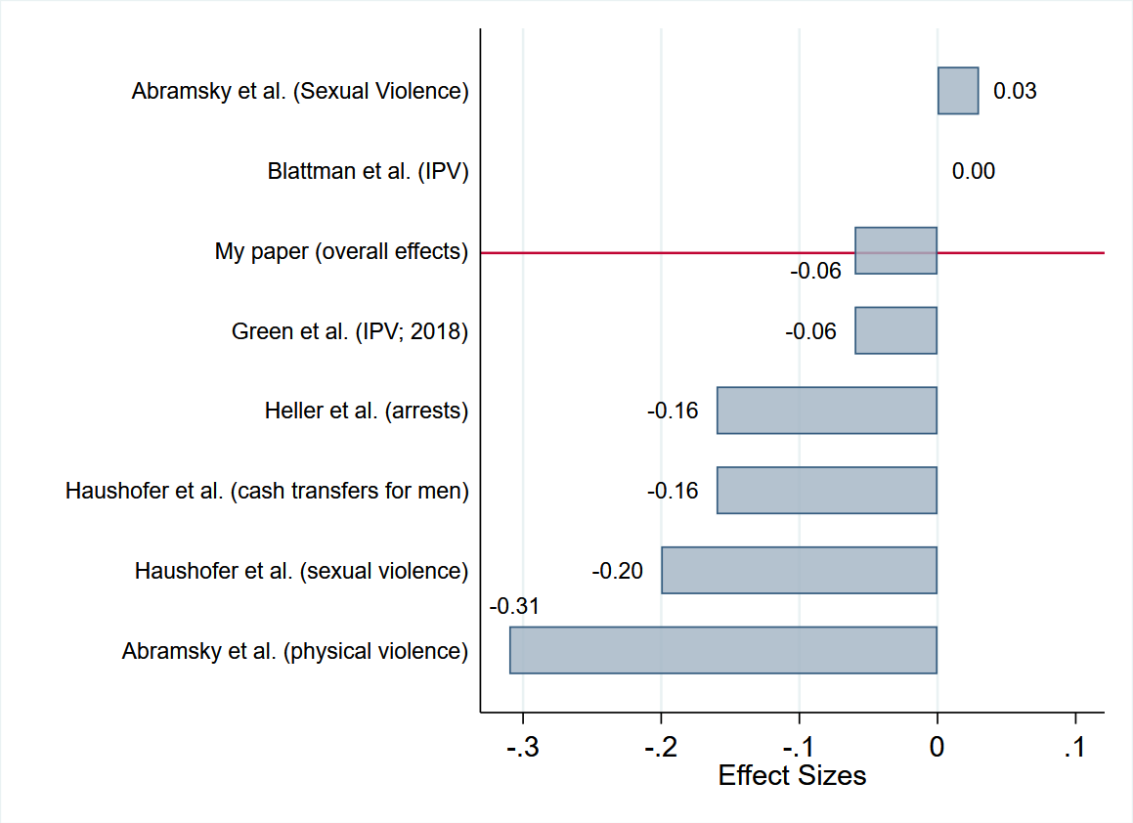


Figure 5: Effect size comparisons with other studies

X axis shows the magnitude of the standardized effect size, and Y axis lists the different studies. The figure shows a comparison of standardized effect sizes with other studies undertaking randomized interventions to deter sexual violence, physical violence and perpetration of other criminal behaviours.

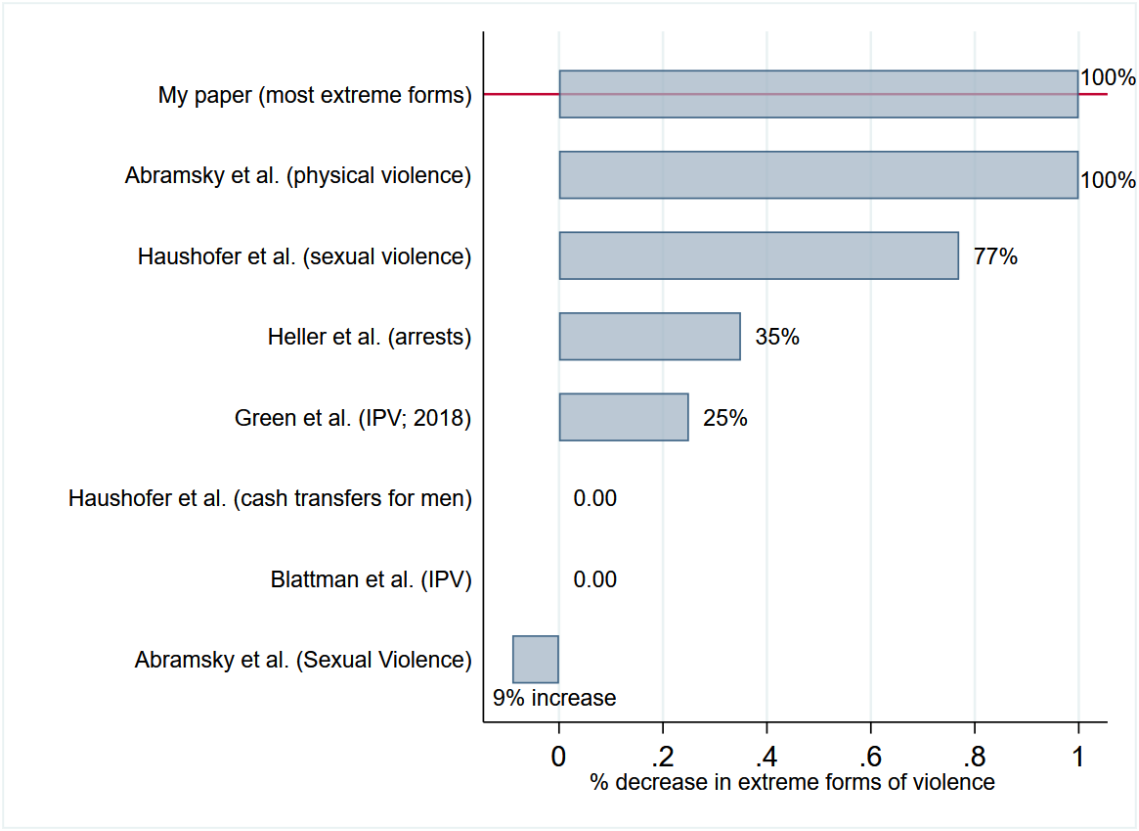


Figure 6: Effect size comparisons with other studies for extreme forms
X axis shows the effect size of interventions on most extreme forms of violence in % terms, and Y axis lists the different studies. The figure shows a comparison of effects on extreme forms of sexual harassment with other studies in % terms.

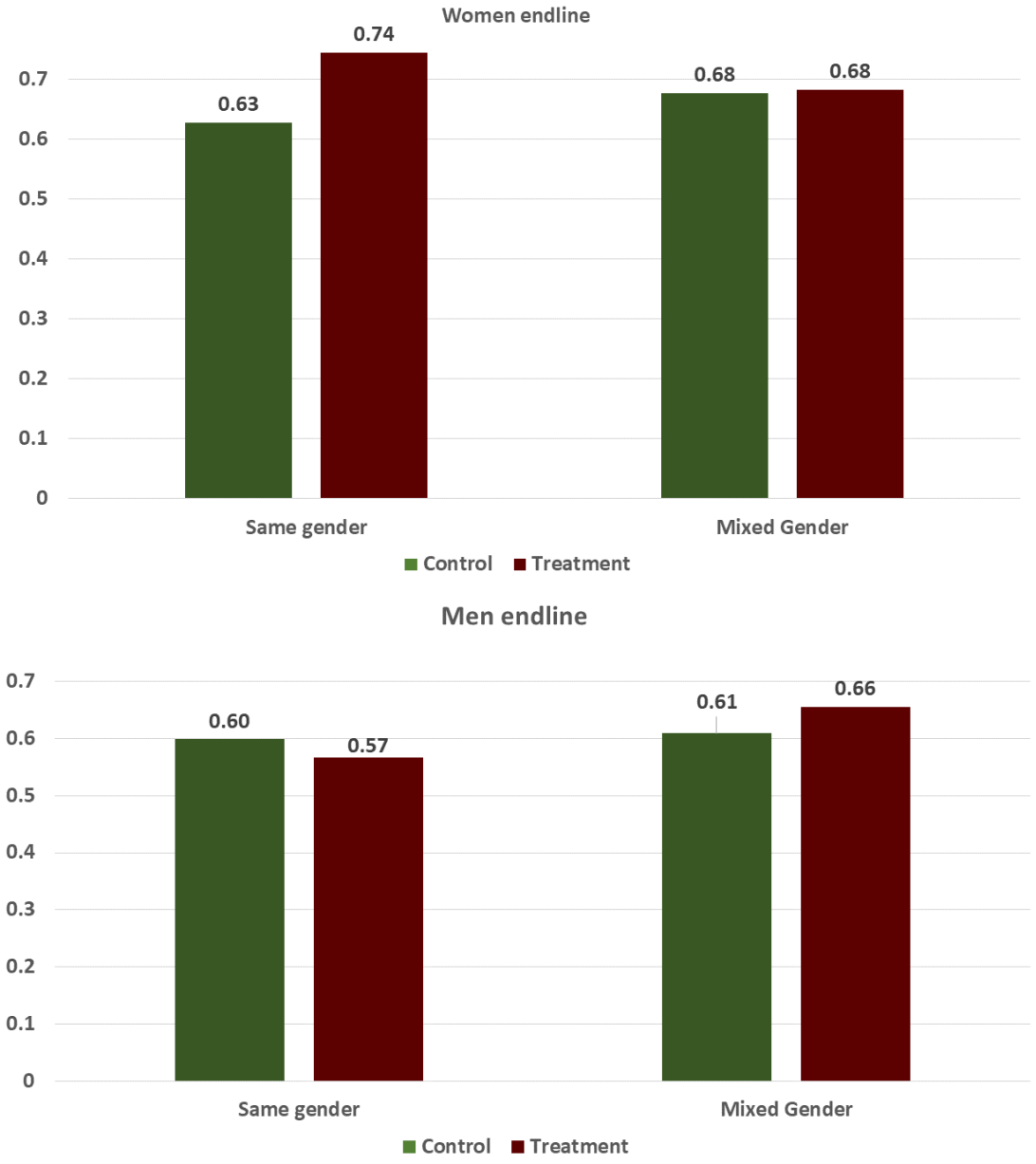


Figure 7: Take up of stick option by gender, treatment and partner's gender
Y axis is the percentage who take up the stick option when paired with a classmate in the lab in the field experiment. Red bars represent treatment groups and green bars are the control group. First panel provides the results for women at endline and the second panel for the men at endline according to the their class' treatment status for the male intervention.

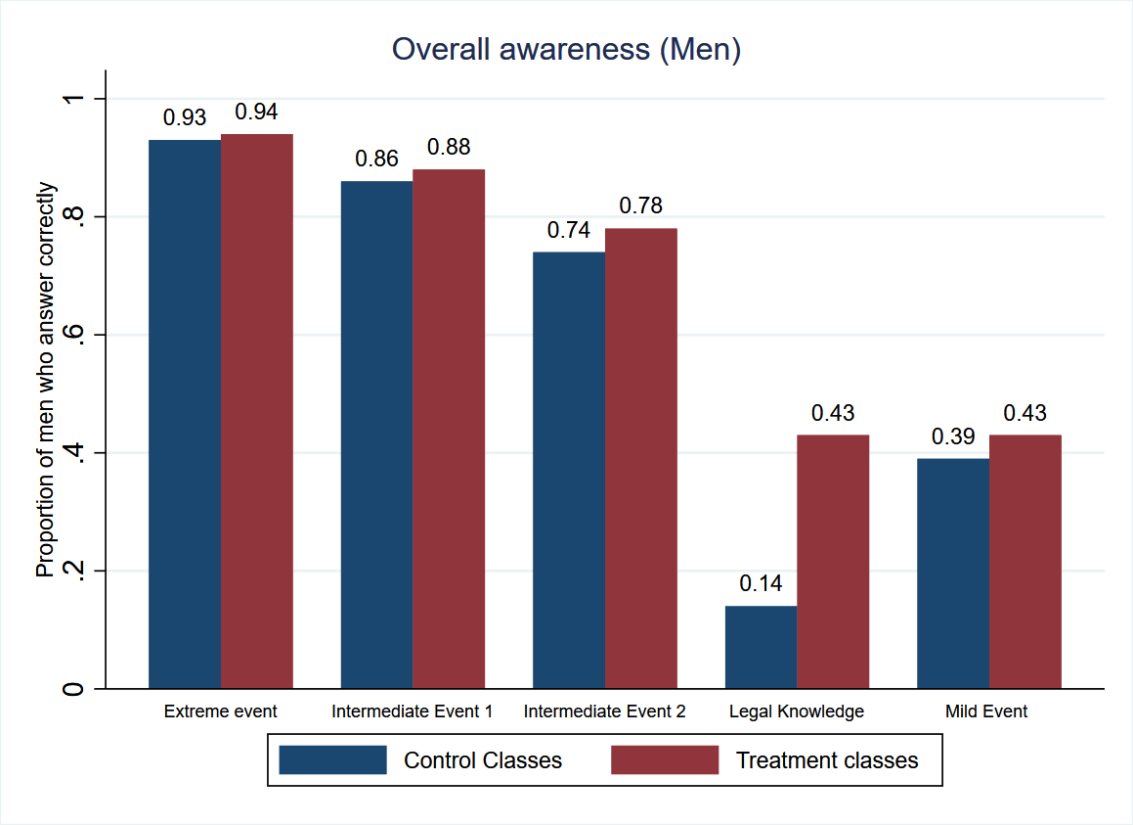


Figure 8: Treatment effects on awareness of men

X-axis lists the different types of events on which I tested men's awareness. Y-axis reports the percentage of men who answered the question correctly. Red bars represent the treatment classes and blue bars are the control classes.

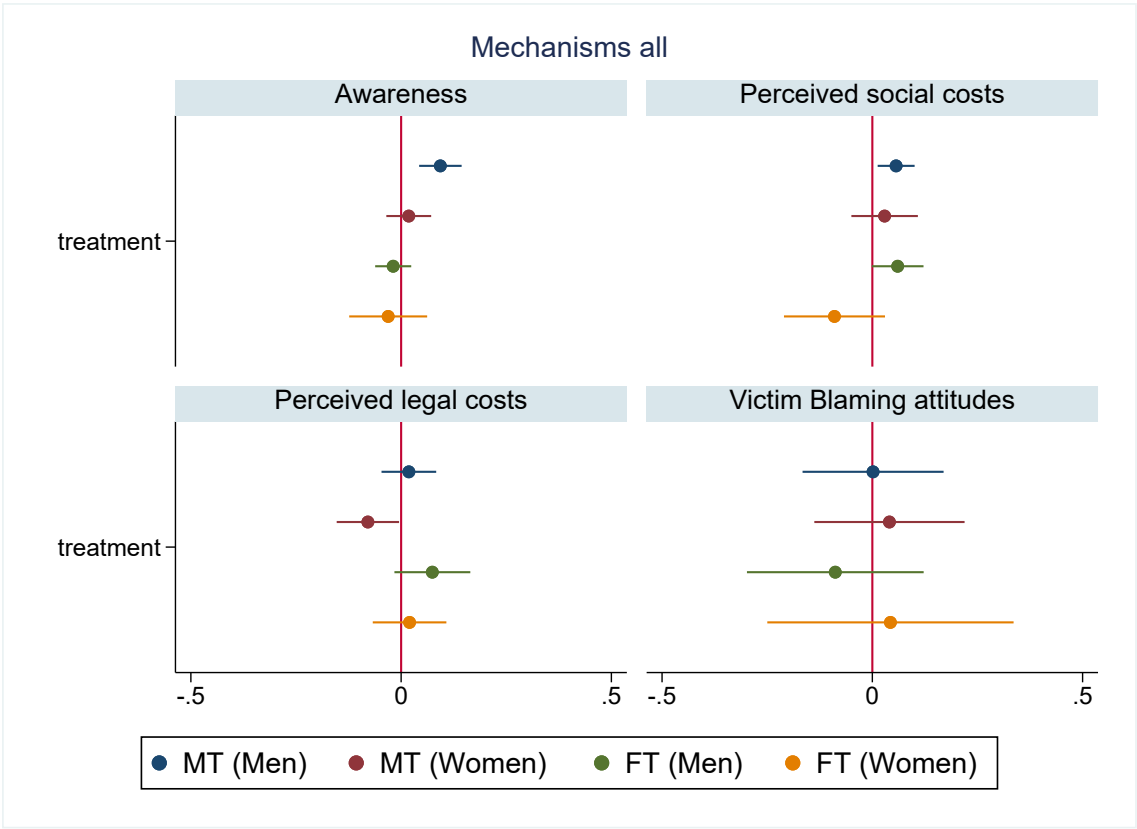


Figure 9: Coefficient plots for effect of female and male intervention on mechanism outcomes for men and women. MT means male intervention and FT means female intervention. The coefficients are created separately for men and women in male and female intervention classes. The coefficients for male intervention are those on the male treatment while coefficients for female intervention are on the class level treatment for women.

Table 1: Sexual harassment prevalence at baseline

Variable	Mean	Std. Dev	Min	Max	N
Low intensity events	0.44	0.49	0	1	1201
Intermediate intensity events	0.47	0.49	0	1	1202
Extreme intensity events	0.16	0.36	0	1	1189

Note: The table reports prevalence rate of sexual harassment of different intensities at the baseline survey with women in the male intervention colleges. Female students were asked about their exposure to different sexual harassment events in the two months prior to the survey. Mild events include sexual remarks, jokes, asking repeatedly out on a date, intermediate events include physical intimidation, stalking, staring, online sexual harassment and extreme events include sexual assault, physical contact without permission like groping, pinching, fondling.

Table 2: Balance Tests for women and men at baseline

Control variable	Treatment Mean	Control Mean	N	p-value
Father education primary	0.05	0.07	2454	0.16
Father education secondary	0.26	0.28	2454	0.28
Father education higher	0.68	0.64	2454	0.16
Mother education primary	0.13	0.14	2413	0.41
Mother education secondary	0.27	0.31	2413	0.16
Mother education higher	0.58	0.54	2413	0.18
Proportion SC/ST/OBC*	0.64	0.62	2675	0.33
Proportion general caste	0.36	0.37	2675	0.34
Proportion other groups	0.01	0.01	2675	0.52
Living in PG/hostel/flat	0.26	0.25	2675	0.89
Living with family	0.74	0.75	2675	0.89
Working mother	0.22	0.23	2902	0.75
Homemaker mother	0.44	0.44	2902	0.93
Whether from Delhi	0.62	0.61	3086	0.64
Number of classes	47	46		
Number of students	1520	1566		
F-stat	1.07			

Note: The table reports mean of baseline characteristics for both men and women in the treatment and control classes. It reports p-values from a regression of the characteristic on the class-level treatment variable. Strata and college FE are included. *SC/ST/OBC represent castes in India. Standard errors are clustered at the class level. p-value for joint test of significance was 0.38.

Table 3: Women's self-reported exposure to sexual harassment perpetrated by men in their class

Sexual Harassment	Same Class Index (1)	Mild events (2)	Intermediate events (3)	Extreme events (4)
No controls, All women				
Male Treatment	-0.0587**	-0.0136	0.0111	-0.0105***
	(0.0286)	(0.0169)	(0.0088)	(0.0036)
With controls, All women				
Male Treatment	-0.0588**	-0.0135	0.0111	-0.0105***
	(0.0282)	(0.0167)	(0.0087)	(0.0035)
Adjusted p values	-	[0.315]	[0.204]	[0.009]***
RI p values	[0.061]*	[0.482]	[0.263]	[0.007]***
N	1255	1195	1165	1165
Control mean (Non-standardized)	-	0.07	0.03	0.01

Note: Reports results from a regression of the dependent variable on class level male intervention dummy variable. *Dependent variable* in **(1)** is an index created using Anderson (2008) method combining questions on different types of sexual harassment perpetrated by men in same class as reported by women in **(2)**, **(3)**, **(4)**. The questions asked female respondent in **(2)** whether they faced any mild event like sexual remarks, jokes, asking repeatedly out on a date from men in their own class, in **(3)** whether they faced intermediate events like physical intimidation, stalking, staring, online sexual harassment from men in their own class and in **(4)** whether they faced extreme events like sexual assault, physical contact without permission like groping, pinching, fondling from men in their own class. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. $PDSLASSO$ is used for selecting controls. Randomization inference p values are reported in square brackets using 1000 repetitions. Adjusted gives p-values are p-values after correcting for multiple hypothesis testing using Benjamini and Hochberg procedure. *p<0.1, **p<0.05 and ***p<0.01. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table 4: Effects of the male intervention on opposite sex relationships

		Survey Measures		Lab-in-the-Field	
	Opposite sex relationship index	Dating (same class)	Opposite sex friends	Switch away (same sex)	Stick to (opposite sex)
	(1)	(2)	(3)	(4)	(5)
All Men					
Male Treatment	0.0423	-0.0103	-0.0061	0.0224	0.0183
	(0.0289)	(0.0071)	(0.0162)	(0.0485)	(0.0484)
N	1895	1539	1810	838	531
Control Mean	0.00	0.02	0.11	0.40	0.61
All women					
Male Treatment	-0.1346***	-0.0129*	-0.0126	-0.1412***	0.0002
	(0.033)	(0.006)	(0.208)	(0.028)	(0.022)
N	1381	1144	1354	555	525
Adjusted p-values	-	[0.08]*	[0.55]	[0.01]**	[0.79]
RI p-value	[0.001]***	[0.11]	[0.60]	[0.03]**	[0.99]
Control mean	-	0.02	0.15	0.37	0.68

Note: Reports results from a regression of dependent variable for men in panel A and women in panel B on the class level intervention for men. Dependent variable in (1) is an index using Anderson (2008) created from a combination of dependent variables in columns 2, 3, 4 and 5. In column 2 the dependent variable is a dummy variable which asked men and women whether they were dating anyone in their own class or not, in column 3 is proportion of opposite gender friends from same class reported by the men and women, in columns 4 it is whether the student switches from same gender partner from their own class or not and in column 5 it is whether the student sticks with the opposite gender partner from their own class or not. Note the number of observations for columns 4 and 5 are less because the lab in field was a between subjects design. Values are thus imputed using KLK method for those who were not in a particular group. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Randomization inference p values are reported in square brackets. *PDSLASSO* is used for selecting controls. P-values adjusted for multiple hypothesis are reported as BH adjusted p-values (Benjamini and Hochberg, 1995). Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table 5: Effect of male sexual harassment awareness training on men and women

	Overall	Situation 1	Situation 2	Situation 3	Situation 4	ICC
	(1)	(2)	(3)	(4)	(5)	(6)
All men						
Male	0.0933***	-0.0044	0.0203	0.0448*	0.0449**	0.1496***
Treatment	(0.0258)	(0.0138)	(0.0182)	(0.0264)	(0.0194)	(0.0227)
N	1904	1624	1580	1423	1904	1904
Control Mean	0.00	0.93	0.86	0.39	0.74	0.14
All women						
Male	0.0404	0.0240*	-0.0056	-0.0694***	-0.0064	0.0630**
Treatment	(0.0355)	(0.0124)	(0.0164)	(0.0261)	(0.0215)	(0.0317)
N	1385	1246	1224	1102	1385	1385
Control Mean	0.00	0.93	0.94	0.50	0.79	0.20

Note: Regression results from estimating equations for dependent variables in columns on class-level intervention for men. Columns 2, 3 and 4 asked men and women to recognize three hypothetical situations as sexual harassment. Dependent variable in Column 2 is a dummy equal to 1 if respondent answers correctly to the question asking about a dating scenario which was not sexual harassment, in column 3 is a dummy equal to 1 if respondent recognizes a hypothetical physical sexual harassment scenario and is a dummy equal to 1 in Column 4 if the respondent correctly identifies a hypothetical ambiguous situation that was sexual harassment. Dependent variable in Column 5 is a dummy equal to 1 if the respondent correctly identifies legally acceptable courtship behaviours from a list of both sexually harassing and non sexually harassing behaviours and in Column 6 is a dummy equal to 1 if the respondent correctly identifies the formal legal complaints committee of their college against sexual harassment. Column 1 is a weighted index of columns 2, 3,4,5 and 6 using Anderson (2008) method. *PDSLASSO* is used for selecting controls. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table 6: Mechanism outcomes

	Perceived social disapproval index (1)	Perceived legal costs index (2)	Victim Blaming Attitudes (3)	Proportion sign-up for volunteership (4)
All men				
Male	0.056**	0.0184	0.001	-0.002
Treatment	(0.0236)	(0.0332)	(0.086)	(0.0141)
N	1904	1887	1851	93
Control Mean	N(0,1)	N(0,1)	0.022	0.02
All women				
Male	0.0157	-0.0790**	0.039	0.006
Treatment	(0.0427)	(0.0379)	(0.093)	(0.0195)
N	1385	1379	1347	93
Control Mean	N(0,1)	N(0,1)	0.021	0.04

Note: Regression results from estimating equations for dependent variables in columns on class-level intervention for men. Dependent variable in column **(1)** is an index created using questions on perception of social disapproval from other classmates, in **(2)** is an index created using questions on perception of legal support for victims, in **(3)** is the coefficient on the interaction between treatment and list treatment that gives the differential effect of the male treatment on number of statements that men (in upper panel) and women (in the lower panel) agree with and in **(4)** measures proportion of students who signed up for volunteering for anti sexual harassment organizations. This variable was collected using google forms floated in all classes and not to each student personally giving 93 classes as the sample sizes. PDSLASSO method was used to select controls. Strata and college fixed effects are included in the regressions. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table 7: Effects of female intervention on sexual harassment

Sexual Harassment	Sexual harassment from same class index	Mild events	Intermediate events	Extreme events
	(1)	(2)	(3)	(4)
Female Treatment β_1	0.0728	0.0247	-0.0009	0.0144
	(0.0994)	(0.0334)	(0.0251)	(0.0229)
High intensity β_2	0.0287	0.0154	0.0429	-0.0128
	(0.0833)	(0.0303)	(0.0350)	(0.0151)
High intensity	-0.1789	-0.0421	-0.0372	-0.0314
$\times$ Female Treatment β_3	(0.1169)	(0.0421)	(0.0395)	(0.0272)
N	563	522	517	554
$\beta_3 + \beta_2 = \beta_1$	0.2419	0.4247	0.8485	0.1735
Control mean	-	0.07	0.05	0.02

Note: Regression results from estimating equations for dependent variables on individual level treatement for women, class level treatment and the interaction. Dependent variables are the same as in table 4. Clustered standard errors are in parenthesis and strata fixed effects are included in all columns. *PDSLASSO* is used for selecting controls. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.001.

Table 8: Effects of female intervention on opposite gender relationships

		Survey Measures		Lab in Field	
	Opposite gender index **(1)**	Dating (same class) (2)	Opposite sex friends (3)	Switch away (same sex) (4)	Stick to (opposite sex) (5)
Female Treatment β_1	-0.0902	-0.0440**	-0.0231	-0.0214	0.0520
	(0.0956)	(0.0171)	(0.0266)	(0.0956)	(0.0816)
High intensity β_2	0.0586	0.0361	-0.0561*	0.0574	-0.0258
	(0.0822)	(0.0343)	(0.0324)	(0.0528)	(0.0649)
High intensity $\times$	-0.1464	-0.0215	0.0160	-0.1412	-0.1097
Female Treatment β_3	(0.1192)	(0.0377)	(0.0435)	(0.1090)	(0.1086)
	595	474	557	167	210
$\beta_3 + \beta_2 = \beta_1$	0.988	0.033	0.741	0.730	0.210
Control mean	0.00	0.035	0.230	0.91	0.14

Note: The table provides estimates from regression of each dependent variable on individual treatment status of the woman, her class's treatment status and the interaction of the two. Column **1** is an index using Anderson method that combines answers to each variable in columns **2** to **5**. Dependent variable is standardization, in **2**, of response to a dummy equal to 1 if the respondent answers in affirmative to holding a romantic relationship with someone of their own class, in **3**, of the proportion of opposite gender friends in same class, in **4**, of dummy equal 1 if switch when paired with same gender partner in the game and in **5** of dummy equal 1 if stick with opposite gender partner in the game. Baseline socio-economic controls are included as well. A missing flag was included for the control if it was missing at the baseline. Strata fixed effects are included and standard errors are clustered at the class level. *p<0.1, **p<0.05 and ***p<0.01.

Table 9: Long run effects of the male intervention on self reported sexual harassment and romantic relationships from second endline only.

	Same class index	Mild events	Intermediate events	Extreme events
	(1)	(2)	(3)	(4)
Male treatment	-0.084	-0.020	-0.023*	-0.008
	(0.058)	(0.020)	(0.014)	(0.009)
N	730	682	696	697
RI p-value	[0.24]	[0.37]	[0.15]	[0.44]
Control Mean	0.00	0.052	0.040	0.013

	Dating in same class	Dating outside the college
Male Treatment	-0.0217**	0.0747**
	(0.010)	(0.031)
N	632	632
RI p-value	[0.046]	[0.035]
Control mean	0.035	0.191

Note: Reports results from a fixed effects regression of the dependent variable on class level male intervention dummy variable using long run surveys with same sample of students. Panel A: *Dependent variable* in **(1)** is dummy variable equals to 1 when a woman reports suffering from mild forms of sexual harassment, in **(2)** from intermediate forms of sexual harassment, and in **(3)** from extreme forms of sexual harassment. The questions definitions are same as in the short run results table. In lower panel, dating in same class is dummy equal 1 if the female student reports being in a relationship with a male student from own class and with someone outside the college. Clustered standard errors are in parenthesis and strata fixed effects and round fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table 10: Long run effects of the male intervention on self reported sexual harassment and romantic relationships with men.

	Mild events (1)	Intermediate events (2)	Extreme events (3)
All women			
Male Treatment	-0.0037	0.0020	-0.0051**
	(0.0089)	(0.0055)	(0.0023)
N	4349	4325	4538
Control mean (Non-standardized)	0.07	0.03	0.01
	Dating in same class	Dating outside the college	
Male Treatment	-0.0173**	0.0613**	
	(0.0072)	(0.0204)	
Control mean	0.035	0.191	
N	1936	1936	

Note: Reports results from a fixed effects regression of the dependent variable on class level male intervention dummy variable using long run surveys with same sample of students. Panel A: *Dependent variable* in **(1)** is dummy variable equals to 1 when a woman reports suffering from mild forms of sexual harassment, in **(2)** from intermediate forms of sexual harassment, and in **(3)** from extreme forms of sexual harassment. The questions definitions are same as in the short run results table. In lower panel, dating in same class is dummy equal 1 if the female student reports being in a relationship with a male student from own class and with someone outside the college. Clustered standard errors are in parenthesis and strata fixed effects and round fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

A Appendix A

A.1 Sexual Harassment outcomes collected from Sexual Harassment Experiences Questionnaire (SEQ)

The specific questions that were asked are as below:

1. Did anyone comment flirtatiously, make direct or indirect remarks/jokes of sexually suggestive or sexist in nature that made you feel uncomfortable?
2. Did anyone try to make unwanted attempts to establish a dating (repeatedly asking you out despite you showing no interest or saying no), romantic or sexual relationship with you despite you trying to discourage it?
3. Did anyone try to get too close to you/try to invade your physical space/brushing against you/cornering you physically in an intimidating, and uncomfortable manner?
4. Did anyone try to watch you, follow you from a distance, stare at you repeatedly making you uncomfortable?
5. Did anyone try to use or display sexual/inappropriate/suggestive material or post vulgar/pornographic/ offensive pictures on messages/email Whatsapp, made some sexual remark or rumours about you?
6. Did anyone try to or attempt to create unwelcome physical contact like pinching you, touching you, groping you or fondling you. (Touching you in areas like thighs, arms, private parts, waist, back, breasts, or your hips) without your permission or consent ?
7. Was anyone able to or attempt to fondle, kiss, or rub against private areas of your body, tried to remove your clothes, or put/insert something into your private body parts without your consent?

Items in 1), 5) represent Gender harassment component of SEQ. Items in 2, 3, 4, 6, 7 are part of unwanted sexual harassment. If a student reports that any of this was done by an administrative authority in college, then I will refer it as sexual coercion. For the purposes of this paper, I asked grouped 1 and 2 for mild, 3,4 and 5 for intermediate and 6 and 7 for extreme forms of harassment. For each of these questions above, I also asked students about the broad category of the perpetrator (someone in their class, someone in same college but not in same class, stranger, administrative member of college, some near home, other, I prefer not to answer this question)

For recall period, the length of the period differed according to when the intervention was done for one college. For majority of the colleges we asked for preceding two months (colleges B, C and D) at baseline. For endline outcomes though we asked for preceding three months.

A.2 Hypothetical sexual harassment scenarios for measuring awareness

Men and women were asked whether they think that the three situations below were sexual harassment. They could answer yes, no and I dont prefer to answer.

1. Harish asked Yashika out on a date. She said yes and went out on a date with him. He asked her to go out with him again but she said no without giving him a reason. Harish got upset about it and asked her why she refused. Yashika told him that she did not think it was fun. Harish agreed and did not ask Yashika out again. Do you think Harish sexually harassed Yashika by calling her to enquire again?

2. Naina and Rahul went out for drinks on a date. Rahul asked Naina if she would like to go dance with him. Naina did join him for the dance. He started touching her physically during the dance but she thought it was because there was not enough space in the dance floor. So she started to dance a bit further from him but he would still end up coming close to her. Do you think Rahul was sexually harassing Naina?

3. Ramit, Arun and Ankur were sitting in the class making some sexual jokes amongst each other. The jokes were not pointed at anyone though. Rita and Smriti who were sitting in the same room having their lunch but could clearly hear what the guys were talking about. Both of them however felt embarrassed and uncomfortable with their conversation but didnot say anything. Were Ramit, Arun and Ankur sexually harassing Rita and smriti?

There were two more questions I asked students to test their awareness about the legal complaints committee and also their awareness about sexual harassment used during courtship.

1. Suppose a young man likes another young woman. They do not have much common friends. Which of the following behaviours according to you are acceptable ways for him to approach her or get to know her? (Please select as many as you find acceptable)

   (a) 1 Get her Whatsapp number from common groups and write to her atleast 2 or 3 times until she replies

   (b) Can stand outside her classes (alone or with friends) to hint that he likes her through indirect comments

   (c) Can send one of his friends to go and talk to her

   (d) Find out her profile on social media (FB, Instagram, Twitter etc) and drop her messages there until she replies

   (e) Find out where she hangs out (clubs or college societies) to understand her schedule on a usual college day

(f) Find out about her by talking to her friends or classmates so that they can tell her that he is interested.

(g) Directly go and talk to her

2. Which of the following is the internal complaints committee of your college? Options: Disciplinary committee, Department teacher in charge, The women development cell, Internal complaints committee, Student Union, Principal, Administrative office, Gender sensitization committee, Other members of faculty, Other.

For courtship question, I coded the answer as correct if the student did not select options a, b, d and e. For 2, the correct answer is internal complaints committee.

A.3 Statements for the list experiment

Students were asked the following: Out of the 3 statements below, can you please tell us how many you agree with? You do not have to state which ones you agree with, but rather how many of them you agree with? Your answer can only be 1,2 or 3 **(and 4 for list treat group)**.

1. Sexual harassment of women by men is a manifestation of the fact that men and women are taught to stay away from each other in our society.

2. Restrictions on mobility of women by parents is valid in the face of increased sexual violence against them.

3. Both, men sexually harassing women and women sexually harassing men or sexual violence in that order should be punishable by law.

4. **Women who face sexual harassment are usually calling for it.**

List control students received only the first three statements and list treatment students received all 4. The difference in the number of statements agreed to by the two groups gives the victim blaming attitude at the class level.

A.4 Hypothetical sexual harassment situations for measuring attitudes.

Men and women were first shown the following three sexual harassment situations.

1. Situation A: Seema was in a college freshers party , when one of her male seniors touched her back discretely. No one else saw (it was dark) but she knew. This will be termed as sexual harassment by Indian law.

2. Situation B: Arun asked Neha out directly for a date first and she agreed. But she did not find it enjoyable while he liked it. Second time however when he asked again and she said no. Arun did not know why she said no. So he called her again but she did not pick up and he kept calling her multiple times to get an answer. This will be termed as sexual harassment by the Indian law.

3. Situation C: Reema was in the bus frequented by other college students as well. She was sitting with one of her male colleagues. He started watching personal pictures on his phone which where sexual in nature. He kept the phone at a distance but Reema was able to see what he was watching. This will be termed as sexual harassment by the Indian law.

The respondents were then asked the following five sets of questions for each of the situations separately.

1. Please tell us for each of the above situations whether they should be termed as sexual harassment legally?

2. Below please tell us for each situation above, what percentage of men who were present in your class right now you think will agree that these were sexual harassment.

3. Below please tell us for each situation above, what percentage of women who were present in your class right now you think will agree that these were sexual harassment.

4. Below please tell us for each situation above, What percentage of women who were present in your class right now you think will report this incident to college's ICC if it happened with them?

5. Below please tell us for each situation, What percentage of women who were present in your class right now you think will report the situations above to other students/teachers/classmates if this happened with them?

A.5 Measuring perception of social and legal costs to sexual harassment

For measuring perceived social costs to sexual harassment, I asked students the following questions.:

1. What percentage of your classmates you think will be supportive of you in case you hypothetically wanted to make a sexual harassment complaint against someone?

2. What percentage of women who are in your class do you think will report or share with other students/teachers/classmates if they were sexually harassed by someone?

3. Who are the three students you would nominate as someone that students in your class can go to for support or advise for going to ICC for a sexual harassment incident?

I used the proportion of male students from their class that women reported to understand if it affected their perception of male support from the class.

For collecting data on perception of formal costs to sexual harassment, I asked the following:

1. What percentage of women who are in your class do you think Will report to college's ICC if they were sexually harassed by someone ?

2. How much do you trust your college's ICC to solve a student's sexual harassment complaint if approached? The options were: Highly trust them, Trust them, Trust them a little, Do not trust them at all.

I constructed an index for perceived legal and social costs separately which I report the results on.

B Anonymized narrative

" *This happened when we were all on the dance-floor, everyone was very close to each other, when I felt a hand on my buttock. At first I brushed it off thinking it must have been a mistake; but then it happened again. This time I was sure I was not imagining it. I looked behind, and I am not sure, till date, who this person exactly was. He was definitely a college senior though, because there was a group of 3-4 seniors dancing right behind us. What scares me till date is the fact that I have no idea who this person was. I am still in touch with most of these seniors, and there is a chance that I am still in touch with my own harasser. It happens on a daily basis, sexual harassment, we have normalized it. But a senior from one's own college doing it is something that disgusts me, and frightens me at the same time. I wish I could have said something that day, screamed, anything; but I was so unsure about what I felt myself, I could not have possibly done anything about it*"

C Comparative statics for theoretical framework

Here, I highlight the proofs for proposition 1. Recall that $c^* = F^{-1}(\frac{-(1-p)u}{pv(D)})$ and $q = \frac{-(1-p)u}{pv(D)} - \frac{v(D)D}{u-v(D)}$.

1. Comparative statics on D.

   - $\frac{dc^*}{dD} = F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{v'(D)}{(v(D))^2})] \geq 0$
   - $\frac{dq}{dD} = \frac{dc^*}{dD} + \frac{uv(D)+uDv'(D)-(v(D))^2}{(u-v(D))^2)}$
   - Thus if;

   $$F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{v'(D)}{(v(D))^2})] + \frac{uv(D)+uDv'(D)-(v(D))^2}{(u-v(D))^2)} \leq 0$$

   then $\frac{dq}{dD} \leq 0$

   Overall, sexual harassment is given by: $p(1-F(c^*))$.
   Therefore, $\frac{dS}{dc^*} = -pF^{'}(c^*)\frac{dc^*}{dD}$ and since we established $\frac{dc^*}{dD} \geq 0$, thus, $\frac{dS}{dc^*} \leq 0$

   Relationships are given by: $R = q[pF(c^*) + (1-p)]$.
   $\frac{dR}{dD} = \frac{dq}{dD}[pF(c^*) + (1-p)] + qpF'(c^*)\frac{dc^*}{dD}$.
   Relationships fall if :
   $F^{-1'}[\frac{-(1-p)u}{pv(D)}][(\frac{-(1-p)u}{p})(-\frac{1}{(v'(D))^2})] \leq -[\frac{u(v(D)+uDv'(D)-(v(D))^2}{(u-v(D))^2}][\frac{pF(c^*)+1-p}{pF(c^*)+(1-p)+qpF'(c^*)}]$

2. Comparative statics on p.

   - $\frac{dc^*}{dp} = F^{-1'}[\frac{-(1-p)u}{pv(D)}](\frac{u}{v(D)p^2}) \leq 0$ because $v(D) < 0$
   - $\frac{dq}{dp} = \frac{dc^*}{dp}$ and hence $\frac{dq}{dp} < 0$.

   $\frac{dS}{dp} = 1 - F(c^*) - pf(c^*)\frac{dc^*}{dp} > 0$
   $\frac{dR}{dp} = \frac{dq}{dp}[pF(c^*) + (1-p)] + q[F(c^*) + pf(c^*)\frac{dc^*}{dp} - 1]$.
   This can be rewritten as: $\frac{dR}{dp} = \frac{dq}{dp}[pF(c^*) + (1-p)] - q\frac{dS}{dp} < 0$. This proves the result.

D Generalizability

Regarding selection, I collaborated with 3 colleges covering all classes (93 in total for male intervention and 69 classes for a female intervention (discussed in the section 6.2)). These colleges are outside of the top 25 colleges as per a list of college rankings for Delhi. For the sample used in the paper, students who were present on the day of the survey (which was unannounced) formed the baseline sample. Students were free to leave the trainings or surveys if they wanted, but less than 5% of the students did so. The surveys took place during regular college hours which helped to get access to student population that would normally be in attendance. Scaling to a bigger population needs to take into account that students in these colleges might be positively selected on household characteristics, ambition etc. that matters for admissions in an urban area. On attrition, we were able to follow-up 80% of the sample in the end line (82% of the women and 77% of men) without any differential attrition by treatment status. Most of the students that were left uncovered (and were supposed to be traced during the college hours) could not be easily reached after college closures due to the lockdown. Third, on naturalness of the choice task and setting, sexual harassment trainings are mandated by law to be undertaken in educational institutions of the type that I collaborate with on the project. Thus, broadly the setting is similar to a target setting of such educational institutes. There is much less generalizability of the setting to workplaces. However both characterize settings that have repeated interactions between potential perpetrators and victims outside home for subjects. The endline consisted of three types of measures i) survey measures ii) lab-in-the-field experiment iii) list experiment and google form data. Survey questions asked students about their exposure sexual harassment (due to lack of any naturally occurring data on incidences).[42] Lab-in-the-field experiment results were used to understand the patterns in survey measures of inter-personal relationships. List experiments and google form data were designed in a way that the purpose of the questions or the exercise was not clear to address demand effects. While the list experiment was embedded within the survey, the google form was floated via the legal complaints committee. In terms of scaling, non-negotiable feature is that the training for men be done with men only and not with women, the timeline over which effects are measured, trust that the participants have in the safety of their data and presence of atleast one male trainer for men. Further replications are important to understand whether providing training to both men and women together leads to different effects.

[42]It was not feasible to have third part observers since that would change behavior of the students even more

E Attrition and Lee bounds, Female information, Difference in difference

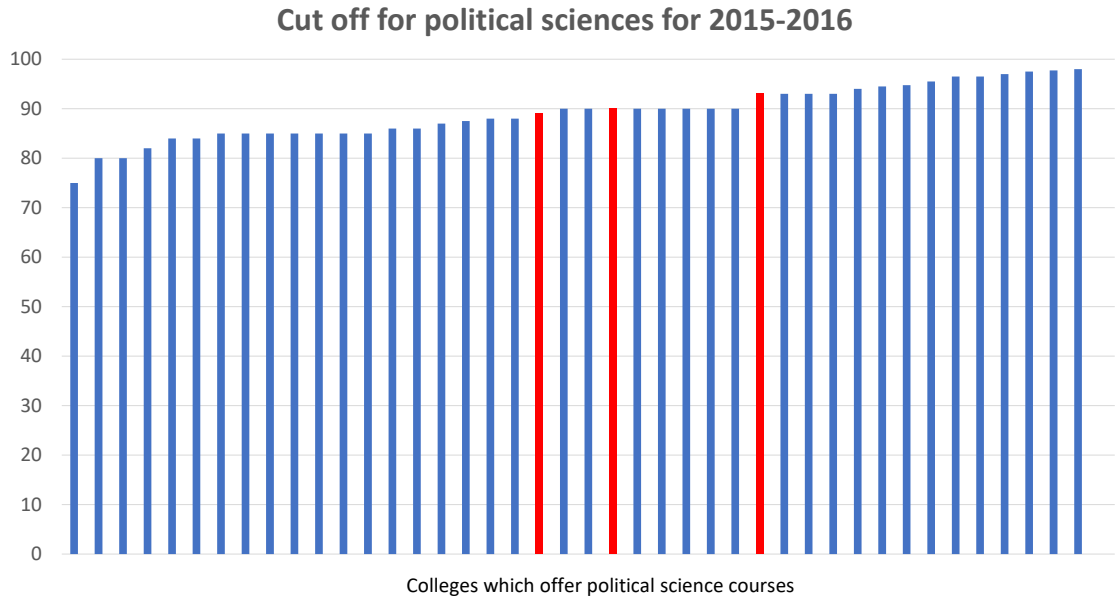


Figure E.1: Perpetrators of sexual harassment as reported by women
The figure above shows the college quality distribution for different colleges in same university as the collaborating college. The colleges in this paper are depicted in red. College quality here is depicted using score cut offs for 2015-2016 academic year for political sciences course that was available for majority of the colleges.

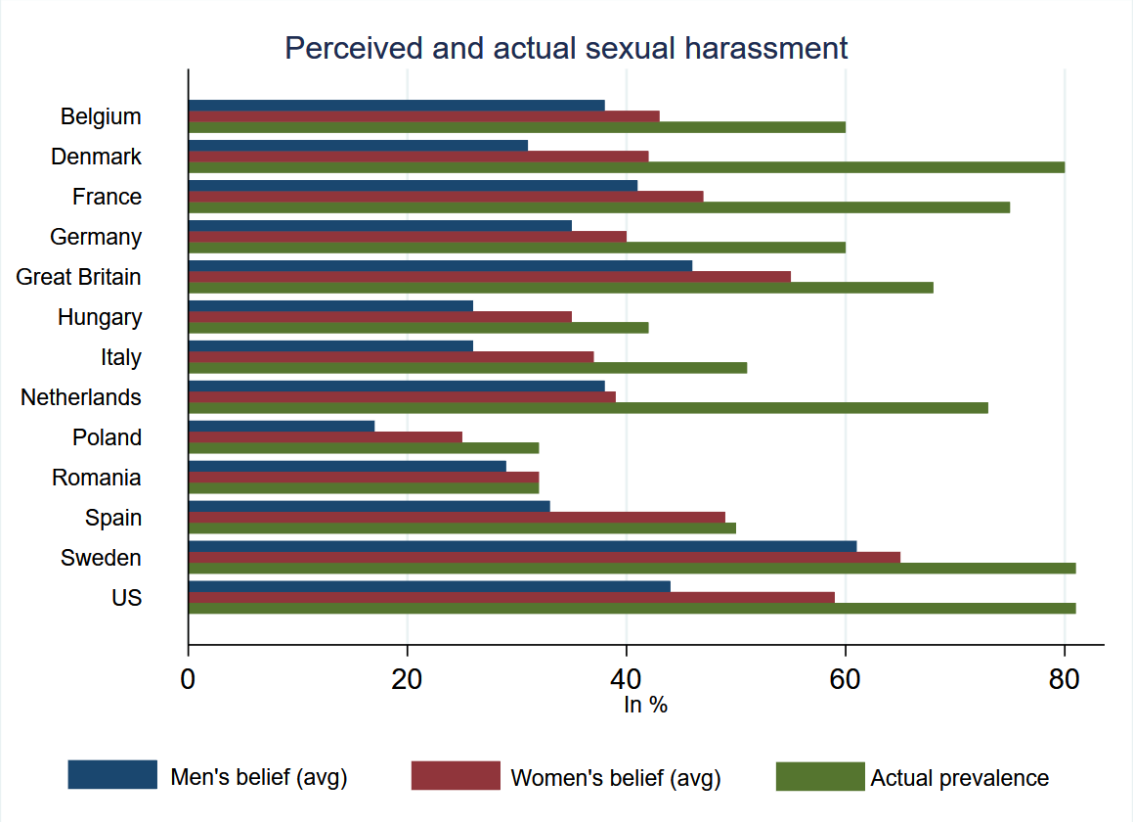


Figure E.2: Average beliefs about prevalence of sexual harassment by country
Figure plots the average beliefs of men and women about the percentage of women they think have been sexually harassed in their own country over a year. X axis in the figure is the average prevalence of sexual harassment for other women. Y axis plots this for men and women for the country referenced. Source: Ipsos, 2018.

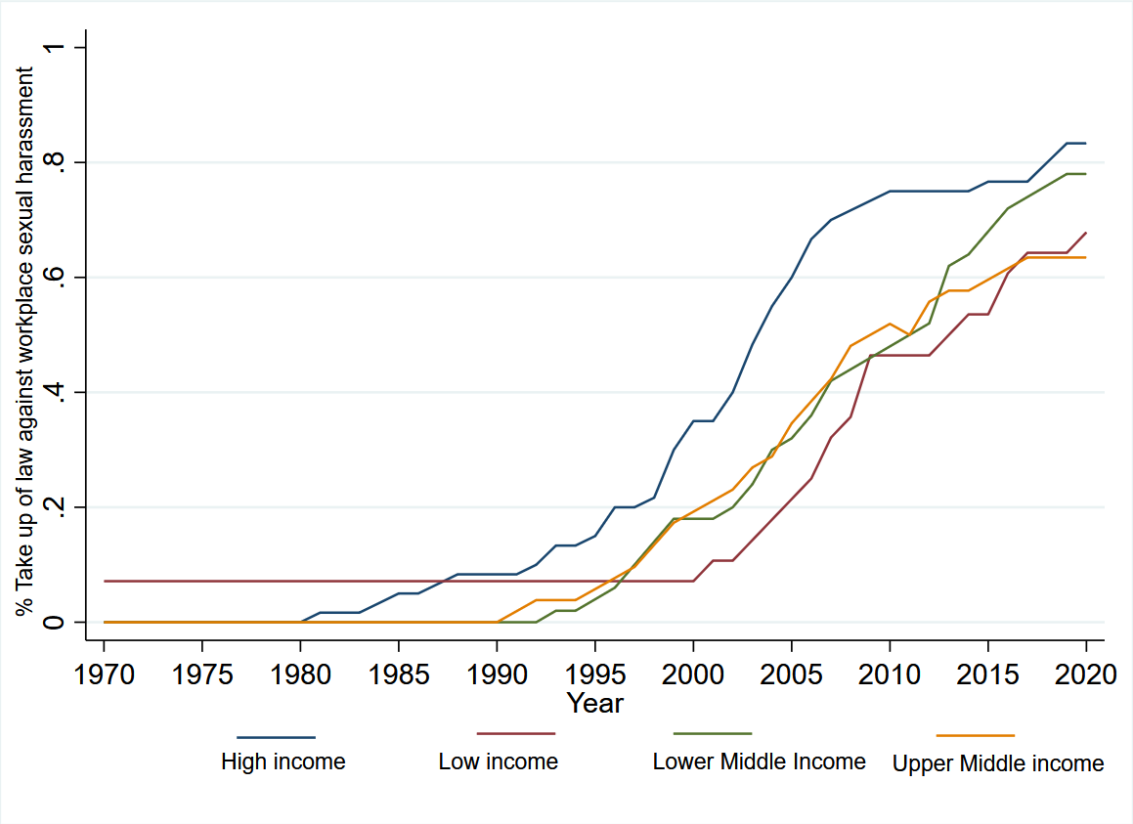


Figure E.3: Take up of laws against sexual harassment (World Bank database)
Figure provides the distribution of beliefs about sexual harassment prevalence for women in their class for men and women in one of the collaborating colleges. X axis is the hypothesized prevalence of sexual harassment in own classroom. Red bars provide the distribution for men's beliefs and green bars for women's beliefs. Both men and women were asked about their beliefs about prevalence of sexual harassment for women only. Actual average prevalence is depicted by the vertical black line.

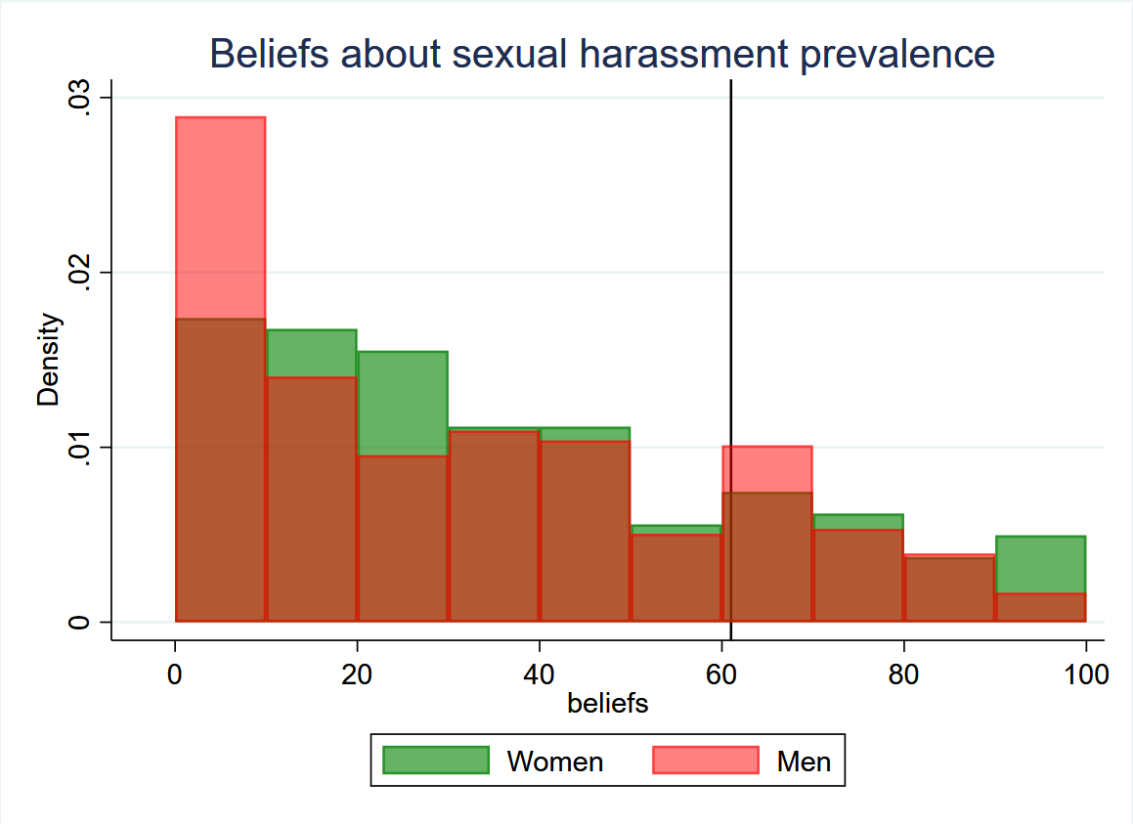


Figure E.4: Distribution of beliefs about prevalence of sexual harassment in college in India
Figure provides the distribution of beliefs about sexual harassment prevalence for women in their class for men and women in one of the collaborating colleges. X axis is the hypothesized prevalence of sexual harassment in own classroom. Red bars provide the distribution for men's beliefs and green bars for women's beliefs. Both men and women were asked about their beliefs about prevalence of sexual harassment for women only. Actual average prevalence is depicted by the vertical black line.

Table E.1: Survey attrition at the endline and treatment

Control variable	Covered in endline	Covered in enline
Male intervention	-0.008	-0.010
	(0.016)	(0.021)
Female		0.060***
		(0.021)
Male intervention × Female		0.006
		(0.032)
N	3075	3075

Note: Dependent variable is a dummy of whether the student was covered in the endline or not. Column (1) reports coefficient from regression on endline coverage on treatment for men and column (2) reports coefficients from regression of attrition on treatment, gender and interaction of the two. Strata and college FE are included. Standard errors are clustered at the college-class level. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.2: Survey attrition at the endline, treatment and baseline controls

Control variable	Covered in endline	p-value
Father education primary	0.109	0.101
Father education secondary	-0.021	0.580
Father education higher	-0.011	0.744
Mother education primary	0.052	0.241
Mother education secondary	0.001	0.988
Mother education higher	-0.028	0.404
Proportion SC/ST/OBC*	-0.028	0.404
Proportion general caste	0.014	0.667
Proportion other groups	0.095	0.787
Living in PG/hostel/flat	-0.033	0.425
Living with family	0.033	0.425
Working mother	0.007	0.815
Homemaker mother	-0.002	0.929
Whether from Delhi	-0.129	0.670
p-value for test of joint significance	-	0.89

Note: The table reports coefficients from interaction term between male intervention and baseline control in a regression of endline coverage on treatment, baseline control and interaction of the two. Strata and college FE are included. Standard errors are clustered at the college-class level. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.3: Effect of female information on recall of sexual harassment using type based questions on sexual harassment

Sexual Harassment	Same Class Index	Mild events	Intermediate events	Extreme events
	(1)	(2)	(3)	(4)
A: All women				
Female treatment	0.0552	0.0320	0.0455*	-0.0028
	(0.0715)	(0.0278)	(0.0248)	(0.0162)
N	735	702	682	684
Control mean (Non-standardized)	-	0.08	0.06	0.03

Note: Reports results from a regression of the dependent variable on individual level female intervention dummy variable. *Dependent variable* in **(1)** is an index created using Anderson 2008 method combining questions on different types of sexual harassment perpetrated by men in same class as reported by women in **(2)**, **(3)**, **(4)**. The questions asked female respondent in **(2)** whether they faced any mild event like sexual remarks, jokes, asking repeatedly out on a date from men in their own class, in **(3)** whether they faced intermediate events like physical intimidation, stalking, staring, online sexual harassment from men in their own class and in **(4)** whether they faced extreme events like sexual assault, physical contact without permission like groping, pinching, fondling from men in their own class. Clustered standard errors are in parenthesis and class fixed effects are included. *p<0.1, **p<0.05 and ***p<0.01. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.4: Short run effect on recall of sexual harassment exposure to self

Dependent variable	Last two months	Sexual Harassment Index (Environment based)	Sexual Harassment Index (Adapted SEQ)
	(1)	(2)	(3)
With class FE			
Women Treatment	0.125***	0.132**	-0.047
	(0.043)	(0.063)	(0.055)
With class FE and controls			
Treatment	0.125***	0.133**	-0.042
	(0.041)	(0.061)	(0.051)
N	690	642	631
Control Mean	0.37	-	-

Note: The table specifies regression coefficient in a regression of the variables above on female information status. This consists of all women who were covered in the baseline survey and for female intervention. Robust standard errors are reported. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.5: Balance Tests for women for female intervention (individual level)

Control variable	β_1	Control Mean	N	p-value
Father education primary	-0.001	0.11	734	0.98
Father education secondary	-0.019	0.27	734	0.60
Father education higher	-0.020	0.53	734	0.60
Mother education primary	-0.041	0.20	728	0.16
Mother education secondary	-0.048	0.32	728	0.21
Mother education higher	0.058	0.36	728	0.12
Proportion SC/ST/OBC*	-0.030	0.38	758	0.43
Proportion general caste	0.040	0.57	758	0.29
Proportion other groups	-0.009	0.01	758	0.33
Proportion Hindu	0.020	0.87	739	0.43
Proportion Muslim	-0.010	0.08	739	0.64
Proportion other religions	-0.010	0.04	739	0.50
Proportion Public transport	-0.052	0.71	743	0.17
Proportion Private paid	0.023	0.10	743	0.38
Proportion self transport	0.029	0.19	743	0.38
Living in PG/hostel/flat	0.025	0.13	703	0.39
Living With family	-0.015	0.77	703	0.13
Working mother	-0.045	0.23	551	0.32

Note: The table specifies regression coefficient in a regression of the variables above on treatment status. This consists of all women who were covered in the baseline survey and for female intervention. β_1 represents coefficient on treatment status of the woman. Robust standard errors are reported. *SC/ST/OBC represent castes in India.

Table E.6: Balance Tests for women for female intervention (class level)

Control variable	Treatment Mean	Control Mean	N	p-value
Father education primary	0.015	0.10	2187	0.19
Father education secondary	-0.007	0.29	2187	0.69
Father education higher	0.001	0.49	2187	0.98
Mother education primary	0.018	0.21	2189	0.20
Mother education secondary	-0.040	0.30	2189	0.01
Mother education higher	0.025	0.33	2189	0.22
Proportion SC/ST/OBC*	0.012	0.39	2262	0.03
Working mother	0.009	0.18	1584	0.65
Whether from Delhi	-0.046	0.54	2254	0.15
Living in PG/hostel/flat	0.030	0.25	2639	0.06
Aims to study after college	0.022	0.72	1976	0.20
Aims to work after college	-0.024	0.24	1974	0.78
Has undertaken job/internship before	-0.004	0.11	2167	0.42
Will sit for job interviews in the future	-0.021	0.46	1975	0.95
Sat for job interviews in the past	-0.004	0.06	2149	0.68
Undertook job through the p-cell of the college	-0.004	0.04	2088	0.61
Number of classes	31	38		
Number of students	972	1347		

Note: The table specifies regression coefficient in a regression of the variables above on treatment status of the class of the student. *SC/ST/OBC represent castes in India. Strata fixed effects are included. Standard errors are clustered at the class level. This includes both men and women in the sample.

Table E.7: Survey attrition at the endline and treatment for female intervention

Control variable	Covered in endline	Covered in enline
Female intervention	-0.006	0.020
	(0.025)	(0.026)
Female		0.176***
		(0.024)
Female intervention × Female		-0.043
		(0.036)
N	2262	2262

Note: Dependent variable is a dummy of whether the student was covered in the endline or not. Female is a dummy equal to 1 if the respondent is a female and 0 if it is a male. Column (1) reports coefficient from regression on endline coverage on treatment for women and column (2) reports coefficients from regression of attrition on treatment, gender and interaction of the two. Strata and college FE are included. Standard errors are clustered at the college-class level. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.8: Survey attrition at the endline, treatment and baseline controls for female intervention

Control variable	Covered in endline	p-value
Father education primary	-0.032	0.567
Father education secondary	0.038	0.321
Father education higher	-0.047	0.196
Mother education primary	-0.045	0.314
Mother education secondary	-0.059	0.148
Mother education higher	0.015	0.718
Proportion SC/ST/OBC*	-0.082	0.053
Whether from Delhi	-0.073	0.045
Living in PG/hostel/flat	0.045	0.374
Aims to study after college	0.000	0.998
Aims to work after college	-0.025	0.608
Has undertaken job/internship before	-0.035	0.538
Will sit for job interviews in the future	0.011	0.791
Sat for job interviews in the past	-0.061	0.475
Undertook job through the p-cell of the college	-0.012	0.917

Note: The table reports coefficients from interaction term between male intervention and baseline control in a regression of endline coverage on treatment, baseline control and interaction of the two. Strata and college FE are included. Standard errors are clustered at the college-class level. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.9: Alternative samples for constructing sexual harassment index

Sexual Harassment	SH from same class Index (only extreme events sample)	SH from same class Index (All reported sample)	SH from same class Index (Simple addition)
A: No controls			
Male Treatment	-0.0624**	-0.0650**	-0.0665**
	(0.0290)	(0.0304)	(0.0311)
B: With controls			
Male Treatment	-0.0624**	-0.0650**	-0.0665**
	(0.0287)	(0.0300)	(0.0308)
RI p values	[0.061]*	[0.063]*	[0.060]*
N	1165	1105	1105
Control mean	0.00	0.00	0.00

Note: Reports results from a regression of the dependent variable (sexual harassment reported by women) on class level male intervention dummy variable. This is estimated only on the sample of women. Column 1 reports for results for only the women who reported extreme events for sure, 2 reports for those who reported all events and uses weights from this sample while 3 reports for those who reported all events but weighting is done using that from the entire sample of women used in the main table. Dependent variable in each column is constructed using the Anderson method as in main table but samples differ across different columns. *Dependent variable* in **(1)** is an index created using Anderson 2008 method for only those women who reported extreme events, in **(2)** is only those women who reported all events, in **(3)** takes simple addition of components of the index. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. $PDSLASSO$ is used for selecting controls. Randomization inference p values are reported in square brackets using 1000 repetitions. B-H procedure gives p-values correcting for multiple hypothesis testing using Benjamini and Hochberg procedure. *p<0.1, **p<0.05 and ***p<0.01. *p<0.1, **p<0.05 and ***p<0.01.

Table E.10: Sexual harassment from men from a different class

Sexual Harassment	Same Class Index (1)	Mild events (2)	Intermediate events (3)	Extreme events (4)
A: With controls, All women				
Male Treatment	-0.0177	-0.0019	-0.0110	0.0018
	(0.0371)	(0.0120)	(0.0140)	(0.0045)
N	1288	1195	1165	1267
Control mean (Non-standardized)	-	0.066	0.046	0.006

Note: *Dependent variable* in **(1)** is an index created using Anderson 2008 method combining different questions asked to women about their exposure to sexual harassment from men in different class asked in **(2)**, **(3)**, **(4)**. The questions asked female respondent in **(2)** whether they faced a mild event like sexual remarks, jokes, asking repeatedly out on a date, in **(3)** whether they faced intermediate events like physical intimidation, stalking, staring, online sexual harassment and in **(4)** whether they faced extreme events like sexual assault, physical contact without permission like groping, pinching, fondling. *p<0.1, **p<0.05 and ***p<0.01. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.11: Sexual harassment from men outside college

Sexual Harassment	Same Class Index	Mild events	Intermediate events	Extreme events
	(1)	(2)	(3)	(4)
A: With controls, All women				
Male Treatment	0.0486	0.0313	0.0055	0.0221*
	(0.0396)	(0.0256)	(0.0249)	(0.0134)
N	1288	1195	1165	1267
Control mean	-	0.27	0.29	0.05

Note: *Dependent variable* in **(1)** is an index created using Anderson 2008 method combining different questions asked to women about their exposure to sexual harassment from men outside college asked in **(2)**, **(3)**, **(4)**. The questions asked female respondent in **(2)** whether they faced a mild event like sexual remarks, jokes, asking repeatedly out on a date, in **(3)** whether they faced intermediate events like physical intimidation, stalking, staring, online sexual harassment and in **(4)** whether they faced extreme events like sexual assault, physical contact without permission like groping, pinching, fondling.. *p<0.1, **p<0.05 and ***p<0.01. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.12: Difference in difference estimates from lab-in-the-field experiment for stick-switch decisions

Decision to stick	(1)	(2)	(3)	(4)
Male treatment	-0.020	-0.025	0.127**	0.123**
	(0.049)	(0.048)	(0.052)	(0.051)
Mixed Gender	0.026	0.027	0.082**	0.076*
	(0.037)	(0.036)	(0.041)	(0.040)
Mixed Gender × Male treatment	0.060	0.057	-0.137**	-0.131**
	(0.062)	(0.061)	(0.058)	(0.056)
N	1369	1369	1080	1080
Control Mean	0.60	0.60	0.63	0.63
Controls	-	Yes	-	Yes

Note: Dependent variable is a dummy variable equal to 1 if the respondent chose to stick with their partner to play the stereotypical task and 0 if they decided to do it alone. Columns 1 and 2 are for men's sample, and columns 3 and 4 are for women's sample. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. *PDSLASSO* is used for selecting controls in columns 2 and 4. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.13: Heterogeneity by year of study for opposite gender relationships index

	Opposite gender index (1)	Dating (same class) (2)	Opposite sex friends (3)	Switch away (same sex) (4)	Stick to (opposite sex) (5)
A: All Women					
Male Treatment (β_1)	-0.1341***	-0.0177**	-3.7186	-0.1265	-0.0497
	(0.0423)	(0.0074)	(2.9975)	(0.0795)	(0.0749)
Older cohort (β_2)	0.1907*	0.0007	6.5104	0.0121	0.1971*
	(0.1051)	(0.0077)	(6.7049)	(0.0864)	(0.1152)
Male Treatment ×	-0.0114	0.0080	3.6465	-0.0264	0.0644
Older Cohort (β_3)	(0.0607)	(0.0131)	(4.1747)	(0.1043)	(0.0967)
$\beta_3 + \beta_2 = \beta_1$	0.005***	0.059*	0.130	0.481	0.063*
N	1381	1144	1354	555	525
Control mean	-	0.02	0.15	0.37	0.68

Reports results from a regression of dependent variable for men in panel A and women in panel B on the class level intervention for men. Older cohort is a dummy equal to 1 if the female respondent belonged to years 2 and 3 of study and 0 otherwise. Dependent variable in (1) is an index using Anderson (2008) created from a combination of dependent variables in columns 2, 3, 4 and 5. In column 2 the dependent variable is a dummy variable which asked men and women whether they were dating anyone in their own class or not, in column 3 is proportion of opposite gender friends from same class reported by the men and women, in columns 4 it is whether the student switches from same gender partner from their own class or not and in column 5 it is whether the student sticks with the opposite gender partner from their own class or not. Note the number of observations for columns 4 and 5 are less because the lab in field was a between subjects design. Values are thus imputed using KLK method for those who were not in a particular group. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Randomization inference p values are reported in square brackets. *PDSLASSO* is used for selecting controls. P-values adjusted for multiple hypothesis are reported as BH adjusted p-values (Benjamini and Hochberg, 1995). Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.14: Sexual harassment for single women in the control group.

	N	Mean	SD	Min	Max
Mild events	494	0.07	0.25	0	1
Intermediate events	488	0.03	0.17	0	1
Extreme events	488	0.01	0.09	0	1

Note:The table above provides prevalence of sexual harassment amongst women in control group who report being single in the endline survey.

Table E.15: Men and women's perception of legal costs to sexual harassment

Dependent variable	Perception of legal costs index (1)	Perceived probability of formal reporting (2)	Trust ICC (3)
A: All men			
Male treatment	0.0184	-0.0137	0.0422
	(0.0332)	(0.0203)	(0.0282)
	1887	1881	1432
B: All women			
Male treatment	-0.0790**	-0.0263	-0.0722**
	(0.0379)	(0.0261)	(0.0331)
N	1379	1377	1067
Control Mean	0.00	0.38	0.51

Note: Regression results from estimating equation regressing dependent variable for men in panel A and for women in panel B on class level intervention for men. Dependent variable in (2) is the perceived probability of formal reporting to ICC by women if sexually harassed (standardized). Dependent variable in (3) *Trust ICC* is answer to the question: "How much do you trust ICC to look into complaints in your college?" (elicited using likert scale) and is standardized as well. Both variables were then combined into a weighted index for column (1) using Anderson (2008). Strata fixed effects are included in all specifications. Clustered standard errors are in parenthesis. *PDSLASSO* is used for selecting controls. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.16: Men and women's perception of social costs to sexual harassment

Dependent variable	Perception of social costs index	Perceived probability of informal reporting	Perception of peer support against sexual harassment	Proportion men nominated for class mascot
	(1)	(2)	(3)	(4)
A: All men				
Male treatment	0.056**	0.020**	0.0286*	-0.0178
	(0.0236)	(0.0103)	(0.0160)	(0.0281)
	1904	1880	1904	1452
Control Mean	0.00	0.36	0.56	0.73
B: All women				
Male treatment	0.0157	0.0240**	0.0221	-0.0387
	(0.0427)	(0.0114)	(0.0208)	(0.0245)
N	1385	1376	1385	1129
Control Mean	0.00	0.38	0.51	0.24

Note: Regression result from estimating equations for the dependent variables, for men in panel A and for women in panel B, on class level intervention for men. Dependent variable in (2) is the perceived probability of informal reporting to friends/classmates by women if sexually harassed, in (3) is *Perception of peer support*: "What percentage of your classmates do you think will support you in case you were to ever seek help or complaint against sexual harassment in college" and in (4) is the proportion of male nominees out of all nominees by the respondent for the position of a class mascot for seeking support after a sexual harassment incident. All variables were then combined into a weighted index for column (1) using Anderson (2008). Strata fixed effects are included in all specifications. Clustered standard errors are in parenthesis. *PDSLASSO* is used for selecting controls. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.17: Men's beliefs about reporting for women to classmates and ICC

What would a woman from your class do in hypothetical sexual harassment situations?	Situation 1 (1)	Situation 2 (2)	Situation 3 (3)
She will report formally			
Male treatment	0.0149	0.0105	-0.0132
	(0.0153)	(0.0147)	(0.0125)
N	1302	1302	1302
Control Mean	0.40	0.28	0.30
She will report informally			
Male treatment	0.0485**	0.0420**	0.0288*
	(0.0202)	(0.0212)	(0.0174)
N	1310	1310	1310
Control Mean	0.45	0.36	0.36

Note: Regression results from estimating equations for dependent variables on class level intervention for men. Columns 2, 3 and 4 asked students whether they think that women in their class are going to report each hypothetical incident formally or informally if it happened to the women. This shows results for the entire endline sample of men. *PDSLASSO* is used for selecting controls in Panel C. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.18: Treatment effect on men's approach behaviour towards women

Variable	Approached by man in same class
	(1)
A: All women	
Male treatment	-0.0224
	(0.036)
N	589
Control mean	0.23

Dependent variable is a dummy which is 1 if woman was approached by a man in her own class and 0 otherwise. Strata fixed effects are included, Clustered standard errors are reported. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.19: Women's relationships with men outside the class

	Dating (outside class) **(1)**	Opposite gender friends (outside class) (2)
All women		
Male treatment	**0.06*** **	-0.030
	(0.023)	(0.062)
N	1,146	1,347
Control mean	0.18	0.00

Dpendent variable in column **1** is a dummy equal to 1 if the respondent answers in affirmative to holding a romantic relationship with someone outside their own class, in **2**, of the proportion of opposite gender friends outside their own class. *p<0.1, **p<0.05 and ***p<0.01.

Table E.20: Men's beliefs about other men and women's perceptions

	Situation 1	Situation 2	Situation 3
	(1)	(2)	(3)
Men's beliefs about male classmates			
Male treatment	0.0659***	0.0611***	0.0422**
	(0.0245)	(0.0142)	(0.0169)
N	1310	1310	1310
Control Mean	0.52	0.35	0.38
Men's beliefs about female classmates			
Male treatment	0.0739***	0.0933***	0.0619***
	(0.0250)	(0.0194)	(0.0201)
N	1310	1310	1310
Control Mean	0.56	0.38	0.43

Note: Regression results from estimating equations for dependent variables on class level intervention for men. Dependent variable is percentage of other classmates that the men thought will agree with the law that the corresponding situation should be legally punishable. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.21: Men's beliefs about other men and women's perceptions

	Situation 1	Situation 2	Situation 3
	(1)	(2)	(3)
Women's beliefs about male classmates			
Male treatment	0.0220	0.0119	0.0237
	(0.0238)	(0.0188)	(0.0185)
N	1022	1022	1022
Control Mean	0.59	0.33	0.36
Women's beliefs about female classmates			
Male treatment	0.0283	0.0143	0.0281
	(0.0260)	(0.0249)	(0.0237)
N	1022	1022	1022
Control Mean	0.64	0.43	0.45

Note: Regression results from estimating equations for dependent variables on class level intervention for men. Dependent variable is percentage of other classmates that the women thought will agree with the law that the corresponding situation should be legally punishable. Clustered standard errors are in parenthesis and strata fixed effects are included in all specifications. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.22: Effects of male intervention on sexual harassment from men in any environment

Sexual Harassment	Same Class Index (1)	Mild events (2)	Intermediate events (3)	Extreme events (4)
A: With controls, All women				
Male Treatment	-0.0243	-0.0114	-0.0092	-0.0135
	(0.0388)	(0.0253)	(0.0247)	(0.0194)
N	1385	1385	1385	1385
Control mean (Non-standardized)	-	0.457	0.458	0.152

Note: *Dependent variable* in **(1)** is an index created using Anderson 2008 method combining different questions asked to women about their exposure to sexual harassment from men in different environments asked in **(2)**, **(3)**, **(4)**. The questions asked female respondent in **(2)** whether they faced a mild event like sexual remarks, jokes, asking repeatedly out on a date, in **(3)** whether they faced intermediate events like physical intimidation, stalking, staring, online sexual harassment and in **(4)** whether they faced extreme events like sexual assault, physical contact without permission like groping, pinching, fondling. *p<0.1, **p<0.05 and ***p<0.01. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.

Table E.23: Long run effects of the male intervention on mechanism outcomes of women

	Awareness Index	Perceived social disapproval index	Perceived legal costs index	Victim Blaming Attitudes	Donation for sign-up for NGO
	(1)	(2)	(3)	(4)	(5)
All women					
Male	0.0217	-0.0546	-0.0182	0.080	0.006
Treatment	(0.0377)	(0.0640)	(0.0513)	(0.118)	(0.073)
N	808	806	806	794	794
Control Mean	N(0,1)	N(0,1)	N(0,1)	2.36	0.51

Note: Regression results from estimating equations for dependent variables in columns on class-level intervention for men. Dependent variable in column **(1)** is anindex created using multiple questions related to awareness of women, in **(2)** is an index created using questions on perception of social disapproval from other classmates, in **(3)** is an index created using questions on perception of legal support for victims, in **(4)** is the coefficient on the interaction between treatment and list treatment that gives the differential effect of the male treatment on number of statements that respondents agree with and in **(5)** measures proportion of students who decided to donate to an NGO campaign against rape in treatment versus control classes. PDSLASSO method was used to select controls. Strata and college fixed effects are included in the regressions. Asterisks denote significance: *p<0.1, **p<0.05 and ***p<0.01.