A Thesis Submitted for the Degree of PhD at the University of Warwick

## Permanent WRAP URL:

http://wrap.warwick.ac.uk/166610

## Copyright and reuse:

This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

# Essays in Development Economics and Economics of Gender 

by

## Karmini Sharma

Thesis submitted for the degree of
Doctor of Philosophy in Economics

Department of Economics, University of Warwick
December 2021

## Contents

List of Figures ..... v
List of Tables ..... ii
Acknowledgements ..... v
Declaration ..... vi
Abstract ..... vii
1 Tackling Sexual Harassment: Experimental Evidence from India ..... 1
1.1 Introduction ..... 2
1.2 Context ..... 7
1.3 Theoretical framework ..... 9
1.3.1 Social Environment ..... 9
1.3.2 Timing ..... 11
1.3.3 Equilibrium ..... 11
1.4 Details of intervention ..... 13
1.5 Experimental Design ..... 15
1.5.1 Recruitment into the awareness intervention ..... 16
1.5.2 Randomization ..... 16
1.5.3 Surveys and intervention administration ..... 17
1.5.4 Sample, timeline and balance tests ..... 19
1.5.5 Data ..... 21
1.5.6 Econometric specification ..... 26
1.6 Results ..... 27
1.6.1 Results for the male intervention ..... 27
1.6.2 Comparison with female sexual harassment awareness inter- vention ..... 39
1.7 Discussion and Conclusion ..... 41
2 Demand for Information by Gender: Evidence a Laboratory Ex- periment ..... 1
2.1 Introduction ..... 2
2.2 Experimental Design ..... 7
2.2.1 The Quiz and Prior Beliefs ..... 7
2.2.2 Choices of Information Structures ..... 8
2.2.3 Signal Received and Posterior Beliefs ..... 11
2.2.4 Signal Received and Posterior Beliefs ..... 11
2.2.5 Debriefing ..... 12
2.2.6 Implementation ..... 12
2.3 Results ..... 13
2.3.1 Descriptive Statistics ..... 13
2.3.2 Prior Beliefs ..... 14
2.3.3 Information Structure Choice ..... 14
2.4 Discussion ..... 20
2.4.1 Posterior Beliefs ..... 20
2.4.2 Implications for Efficiency ..... 22
2.4.3 Overall results ..... 22
2.5 Concluding Remarks ..... 24
3 Attribution Bias by Gender: Evidence a Laboratory Experiment ..... 1
3.1 Introduction ..... 2
3.2 Literature Review ..... 3
3.3 Experimental Design ..... 5
3.3.1 The Experiment ..... 5
3.3.2 Gender Information ..... 7
3.4 Experimental Hypotheses ..... 9
3.5 Experimental Results ..... 10
3.5.1 Implementation ..... 10
3.5.2 Agents ..... 11
3.5.3 Principals ..... 12
3.5.4 Econometric Specification ..... 17
3.5.5 Econometric Results ..... 18
3.5.6 Robustness Checks ..... 20
3.6 Possible Threats ..... 21
3.6.1 Gender Information ..... 21
3.6.2 Prior Beliefs by Gender of the Agent ..... 22
3.6.3 Selection of our Sample ..... 22
3.6.4 Agents' Age ..... 22
3.7 Discussion and Conclusion ..... 23
A Chapter 1 Appendix ..... 24
A. 1 Appendix A ..... 24
A.1.1 Sexual Harassment outcomes collected from Sexual Harass- ment Experiences Questionnaire (SEQ) ..... 24
A.1.2 Hypothetical sexual harassment scenarios for measuring aware- ness ..... 25
A.1.3 Statements for the list experiment ..... 26
A.1.4 Hypothetical sexual harassment situations for measuring atti- tudes. ..... 27
A.1.5 Measuring perception of social and legal costs to sexual ha- rassment ..... 28
A. 2 Anonymized narrative ..... 28
A. 3 Comparative statics for theoretical framework ..... 30
A. 4 Generalizability ..... 31
A. 5 Attrition and Lee bounds, Female information, Difference in difference ..... 32
B Chapter 2 Appendix ..... 52
B. 1 Appendix B ..... 52
B.1.1 Screenshots of the Experiment ..... 52
B.1.2 Welcome Page ..... 52
B.1.3 Instructions - the Quiz ..... 53
B.1.4 Comprehension Questionnaire - the Quiz ..... 54
B.1.5 The Quiz ..... 55
B.1.6 Instructions - Belief Elicitation ..... 56
B.1.7 Prior Belief Elicitation ..... 57
B.1.8 Instructions - Information Structure Choice (T1) ..... 58
B.1.9 Comprehension Questions - Information Structure Choice ..... 61
B.1.10 Information Structure Choice ..... 62
B.1.11 Partner's Feedback ..... 63
B.1.12 Posterior Belief ..... 64
B.1.13 Instructions for strategic treatment (T4) ..... 65
B.1.14 Instructions for strategic treatment (T4) continued ..... 66
B.1.15 Instructions for strategic treatment (T4) continued ..... 67
B.1.16 Distribution of questions answered correctly ..... 68
B.1.17 Qualitative answers ..... 69
B.1.18 Efficiency versus equity ..... 70
B. 2 Summary Statistics ..... 71
B.2.1 Demographic questions ..... 72
C Chapter 3 Appendix ..... 80
C. 1 Appendix C: Summary Statistics ..... 80
C.1.1 Mean Payments to Each Party by Realized Outcome ..... 80
C.1.2 Mean Payments to Each Party by Realized Outcome and Gen- der of the Agent ..... 80
C.1.3 Robustness Checks for Principals' Payment Decisions and Beliefs ..... 80

## List of Figures

1.1 Perpetrators of sexual harassment as reported by women ..... 9
1.2 Overall Design ..... 17
1.3 Male training ..... 18
1.4 Female surveys ..... 19
1.5 Effect size comparisons with other studies ..... 30
1.6 Take up of stick option by gender, treatment and partner's gender ..... 34
1.7 Treatment effects on awareness of men ..... 36
1.8 Coefficient plots for effect of female and male intervention on mech- anism outcomes for men and women. ..... 43
2.1 Information Structures in the Experiment ..... 9
2.2 Take Up of Information Structure B by Strategic/Non Strategic En- vironment ..... 19
3.1 Principals' payment decisions by realized outcome ..... 13
3.2 Principals' beliefs by realized outcome ..... 14
3.3 Principals' wages by realized outcome and agents' gender ..... 15
3.4 Principals' beliefs by realized outcome and agents' gender ..... 16
A. 1 Average beliefs about prevalence of sexual harassment by country ..... 33
A. 2 Take up of laws against sexual harassment (World Bank database) ..... 34
A. 3 Distribution of beliefs about prevalence of sexual harassment in col- lege in India ..... 35
B. 1 Screenshot of the Welcome Page Template ..... 52
B. 2 Screenshot of the Instructions Template for the Quiz ..... 53
B. 3 Screenshot of the Comprehension Questionnaire Template for the Quiz ..... 54
B. 4 Screenshot of the Quiz ..... 55
B. 5 Screenshot of the Instructions Template for the Prior Belief Elicitation ..... 56
B. 6 Screenshot of the Prior Belief Elicitation ..... 57
B. 7 Screenshot of the Instructions Template for the Information Structure Choice - Part I ..... 58
B. 8 Screenshot of the Instructions Template for the Information Structure Choice - Part II ..... 59
B. 9 Screenshot of the Instructions Template for the Information Structure Choice - Part III ..... 60
B. 10 Screenshot of the Comprehension Questionnaire Template for the In- formation Structure Choice ..... 61
B. 11 Screenshot of the Information Structure Choice ..... 62
B. 12 Screenshot of the Feedback the Subject Received about her Partner ..... 63
B. 13 Screenshot of the Feedback Received and Elicitation of Posterior Belief ..... 64
B. 14 Screenshot of information about payments for Strategic Treatment ..... 65
B. 15 Screenshot of information about payments for Strategic Treatment ..... 66
B. 16 Screenshot of information about payments for Strategic Treatment ..... 67
B. 17 Distribution of number of questions answered correctly by treatment. ..... 68
B. 18 Qualitative answers to reasons behind choosing less informative feed- back. ..... 69
B. 19 Efficiency versus equity. ..... 70
C. 1 Mean payments to each party by realized outcome ..... 82
C. 2 Mean payments to each party by realized outcome and gender of the agent ..... 82

## List of Tables

1.1 Sexual harassment prevalence at baseline ..... 8
1.2 Balance Tests for women and men at baseline ..... 20
1.3 Women's self-reported exposure to sexual harassment perpetrated by men in their class ..... 29
1.4 Effects of the male intervention on opposite sex relationships ..... 32
1.5 Effect of male sexual harassment awareness training on men and women ..... 37
1.6 Mechanism outcomes ..... 38
1.7 Effects of female intervention on sexual harassment ..... 40
1.8 Effects of female intervention on opposite gender relationships ..... 42
2.1 Main Features of the Treatments ..... 10
2.2 Gender differences in the choice of information structure A by treat- ment ..... 15
2.3 Choice of Feedback Mode B in (T1) ..... 16
2.4 Choice of Information Structure B in (T1)-(T3) ..... 17
2.5 Overall difference in take up of less informative feedback, all treatments ..... 21
3.1 Mean performance by task and gender ..... 11
3.2 Mean beliefs about performance by task and gender ..... 11
3.3 Mean beliefs about realized outcomes and expected principals' payments ..... 12
3.4 Regression results for principal's payments ..... 18
3.5 Regression results for principal's beliefs ..... 19
A. 1 Survey attrition at the endline and treatment ..... 32
A. 2 Survey attrition at the endline, treatment and baseline controls ..... 36
A. 3 Effect of female information on recall of sexual harassment using type based questions on sexual harassment ..... 37
A. 4 Short run effect on recall of sexual harassment exposure to self ..... 38
A. 5 Balance Tests for women for female intervention (individual level) ..... 39
A. 6 Balance Tests for women for female intervention (class level) ..... 40
A. 7 Survey attrition at the endline and treatment for female intervention ..... 41
A. 8 Survey attrition at the endline, treatment and baseline controls for female intervention ..... 42
A. 9 Alternative samples for constructing sexual harassment index ..... 43
A. 10 Sexual harassment from men from a different class ..... 44
A. 11 Sexual harassment from men outside college ..... 44
A. 12 Difference in difference estimates from lab-in-the-field experiment for stick-switch decisions ..... 45
A. 13 Heterogeneity by year of study for opposite gender relationships index ..... 46
A. 14 Sexual harassment for single women in the control group. ..... 47
A. 15 Men and women's perception of legal costs to sexual harassment ..... 47
A. 16 Men and women's perception of social costs to sexual harassment ..... 48
A. 17 Men's beliefs about reporting for women to classmates and ICC ..... 49
A. 18 Treatment effect on men's approach behaviour towards women ..... 49
A. 19 Women's relationships with men outside the class ..... 50
A. 20 Men's beliefs about other men and women's perceptions ..... 50
A. 21 Men's beliefs about other men and women's perceptions ..... 51
A. 22 Effects of male intervention on sexual harassment from men in any environment ..... 51
B. 1 Descriptive statistics ..... 71
B. 2 Descriptive statistics by information structure chosen ..... 73
B. 3 Choice of Feedback Mode B in (T1) with session fixed effects ..... 74
B. 4 Choice of Information Structure B in (T1)-(T3) with session FE ..... 75
B. 5 Overall difference in take up of less informative feedback, all treat- ments with session fixed effects ..... 76
B. 6 Choice of Feedback Mode B in (T1) with standard errors clustered at the session level ..... 77
B. 7 Choice of Information Structure B in (T1)-(T3) with clustered stan- dard errors ..... 78
B. 8 Overall difference in take-up of less informative feedback, all treat- ments with clustered standard errors ..... 79
C. 1 Summary statistics of our sample ..... 81
C. 2 Summary statistics of variables in main econometric specification ..... 81
C. 3 Regression results for principals' payments with session fixed effects ..... 83
C. 4 Regression results for principals' payments with round fixed effects ..... 84
C. 5 Regression results for principals' payments without controls ..... 84
C. 6 Regression results for principals' payments for first ten rounds only ..... 85
C. 7 Regression results for principals' payments after removing the first five rounds ..... 86
C. 8 Regression results for principals' beliefs with session fixed effects ..... 87
C. 9 Regression results for principals' beliefs with round fixed effects ..... 88
C. 10 Regression results for principals' beliefs without controls ..... 88
C. 11 Regression results for principals' beliefs for the first ten rounds only ..... 89
C. 12 Regression results for principals' beliefs after removing the first five rounds ..... 90
C. 13 Regression results for different parts of the cumulative distribution of payments ..... 90

## Acknowledgements

I am grateful for the support I have received throughout my doctoral studies from my family, supervisors, colleagues, and friends; and I would like to thank them all for helping me reach this point.

First and foremost, my deepest gratitude goes to my parents, Pallavi and Gautam. They kept me grounded and inspired throughout the ups and downs in my PhD journey. I thank Pallavi for being my rock throughout these years and beyond. I have truly received a gift in you as my sister.

Secondly, I deeply thank my supervisors Clément Imbert and James Fenske for their professional guidance, support, and mentorship that has brought me to this point in my PhD . Thank you for the countless discussions and meetings for your thoughts throughout these past five years. My professional journey was impossible without you and your vision. I am specially indebted to Clément Imbert for his support, countless discussions, and guidance during field work. It was paramount in making sure I never lost track of the goals. Deep gratitude to you. I also thank Bishnupriya Gupta, Lori Beaman, Debraj Ray, and Christopher Roth for their mentorship, nudges and guidance in the process.

I am thankful for the broader support I had on a personal and professional level from others at different points in my journey especially Felix, Diego, Mariana, Apurav, Yatish, Alessandro, Roberto, Nivi, Valeria, Stefania, Kanika, Prerna and Andy. I thank Pooja and Najma who supported me since the first day on my field work through some of the most challenging moments. A big thank you to Natalie Deven and Maryanne Heafey for support in so many administrative activities during the entire journey.

Finally, I would like to acknowledge generous financial support I have received for my doctoral studies and for the specific research projects that comprise my thesis. This includes funding from the Department of Economics at Warwick, Chancellor's fund, Crime and Violence Initiative at J-PAL and Bill and Melinda Gates Foundation.

## Declaration

I submit this thesis to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy in Economics. I declare that it has not been submitted for a degree at another university. Chapter 1 is solo-authored. Chapter 2 is co-authored with Alessandro Castagnetti (University of Warwick), where we both were equally involved in developing the project design, arranging for funding, implementing the lab experiments and data analysis. This chapter is now at a revise and resubmit stage at the Journal of Economic Behaviour and Organization. Chapter 3 is co-authored with Alessandro Castagnetti (University of Warwick) and James Fenske (University of Warwick). James Fenske facilitated the access to the funding for this project. Alessandro and I were both equally involved in developing the design, implementing the lab experiments and undertaking the data analysis. We wrote the academic paper together and James Fenske also guided us in writing the paper better.

December 2021


#### Abstract

This thesis studies three topics in development economics and economics of gender as summarized below.

Chapter 1 provides results from a randomized controlled trial undertaken in collaboration with colleges in New Delhi to study the effect of sexual harassment awareness training for men on women's self-reported sexual harassment and on relationships between men and women. I find that approximately three to four months after the training, both sexual harassment and opposite sex relationships go down. I show that this is consistent with an increase in men's perceived peer disapproval against sexual harassment in their class.

Chapter 2 studies gender gap in endogenous information seeking about one's own performance. We design a laboratory experiment in which subjects perform a male stereotypical task and then choose to receive feedback from two differentially informative feedback structures. We then introduce variations in the costs and benefits of choosing more informative feedback structure. We find that women seek less information than men but that introducing strategic concerns removes this gender gap.

Chapter 3 tests for attribution bias by gender which means attribution of unexpectedly good outcomes to ability for men and luck for women and vice versa, in case of unexpectedly bad outcomes. We set up a principal-agent framework using a laboratory experiment. Principal's outcomes are affected by both the effort of the agent and a random component. In particular, higher the number of questions answered correctly by the agent, the higher is the expected payoff from a lottery that is assigned to the principal. We then test whether the principal differentially pays according to the gender of the agent. We do not find evidence of attribution bias by gender but find instead that the principals pay lesser to agents of the same gender.


## 1 Tackling Sexual Harassment: Experimental Evidence from India

Sexual harassment imposes substantial economic costs on the victims, yet there is limited evidence on how to effectively deter it. I present experimental evidence on the effects of a sexual harassment awareness training for college students in New Delhi, using a randomized controlled trial. I find that sexual harassment awareness training for men reduces sexual harassment reported by women in their peer groups by 0.06 standard deviations. 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 theoretical framework of signalling, 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. Empirically, 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. 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. Thus, this paper shows that it is possible to engage men for women's empowerment.

### 1.1 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

[^0]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 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

[^1]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 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 attitude in their class. All these results are consistent with the second mechanism. Third, I also measured men's attitudes using a list experiment and a volunteer-ship 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

[^2]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 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.

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); 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, gen-

[^3]der segregation. I contribute by studying how the training affects opposite sex relationships in colleges. I 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); 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 1.2, the theoretical framework in section 1.3, the details of the intervention in section 1.4, the experimental design in section 1.5, the results in section 1.6 and conclude in section 1.7.

### 1.2 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 A.2. 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 A.3. 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 workplaces 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

[^4]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 1.1. Low and intermediate intensity events were highly 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.

Table 1.1: 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.

Next, I show who perpetrated these sexual harassment incidents in figure 1.1. 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.


Figure 1.1: 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.

### 1.3 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)).

### 1.3.1 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 $\operatorname{cost} c_{i}$ if they undertake $g$ where $c_{i} \sim f($.$) over [0, \infty)$. Thus, a $B$ type man is characterized by $\left(t_{i}, c_{i}\right)$ 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 $\mathrm{P}($.$) . Men's utility is characterized by:$

$$
\begin{aligned}
U\left(t_{i}, a_{i}\right)=\underbrace{I\left(W \text { accepts } \mathrm{a}_{\mathrm{i}}\right)}_{\text {Pairing utility }} & -\underbrace{c_{i} I\left(a_{i}=g, t_{i}=B\right)}_{\text {Psychic costs for B types }}+\underbrace{k I\left(a_{i}=g, t_{i}=G\right)}_{\text {Intrinsic utility for } \mathrm{G} \text { types }} \\
& -\underbrace{D P\left(t_{i}=B \mid a_{i}\right)}_{\text {Social disapproval }}
\end{aligned}
$$

$I($.$) is an indicator function that takes a value 1$ 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 $\mathrm{D} . v($.$) is assumed$ 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 $v^{\prime}(D) \geq 0$. 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.

[^5]
### 1.3.2 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.
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: $\operatorname{Pr}\left(t_{i}=B \mid a_{i}\right)$ and $\operatorname{Pr}\left(t_{i}=G \mid a_{i}\right)$.
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.

### 1.3.3 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 $P\left(t_{i}=G \mid a_{i}\right) \geq \frac{-v(D)}{u-v(D)}$ 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 $\left(a_{G} \in\{b\right.$, 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 $P\left(c^{*}\right)$ where $P\left(c^{*}\right)=\operatorname{Pr}\left(t_{i}=B \mid a_{i}\right)$ and $P\left(c^{*}\right):\{b, g\} \rightarrow[0,1]$.
3. Woman's strategy for each action of the man, $a_{w}:\{b, g\} \rightarrow\{$ 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 $g$.
(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

[^6]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, $\operatorname{Pr}\left(t_{i}=G \mid a_{i}=g\right)=\frac{-v(D)}{u-v(D)}$ and $\operatorname{Pr}\left(t_{i}=G \mid a_{i}=b\right)=0$. Thus, the beliefs follow Bayes' rule on the equilibrium path. Thus a fraction $F\left(c^{*}\right)$ 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 $a_{i}=g$. Thus, total prevalence of sexual harassment is given by $\left(1-F\left(c^{*}\right)\right) p$ and total relationships are given by $q\left[F\left(c^{*}\right) p+(1-p)\right]$. Notice that by Bayes' rule, $\frac{p F\left(c^{*}\right)}{p F\left(c^{*}\right)+(1-p)}=\frac{u}{u-v(D)}$. This gives that $c^{*}=F^{-1}\left(\frac{-(1-p) u}{p v(D)}\right)$. Given the belief, $P\left(c^{*}\right)$, 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 $-D=-c^{*}+q-D P\left(c^{*}\right)$ so that the costs and benefits of doing $g$ for him are equalized. This can be rewritten to give $q=F^{-1}\left(\frac{-(1-p) u}{p v(D)}\right)+\frac{v(D) D}{u-v(D)}$. Thus, woman's acceptability of $g$ is dependent on D and $p$.

Proposition 2. 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 A.3.
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 $v^{\prime}(D)>0$ ). Overall effect on relationships is, thus, ambiguous for increase in D even
though sexual harassment decreases. However, I show in appendix A. 3 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: 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.

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.

Finally, I hypothesize that the female intervention changes women's perceptions of $p$. Thus, for women there are now more B types in their environment. From prediction 2, we know that this means that women's acceptance of men's offer of relationships will decrease and relationships ought to fall. However, to study equilibrium effects of this shift, we cannot assume that men can change $c^{*}$ since men are still unaware. Their lack of ability to distinguish between $b$ and $g$ necessarily means either no change in their behaviour or taking no action towards women. This can reduce sexual harassment but also relationships due to both men and women's behaviour.

### 1.4 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.1. 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 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.

### 1.5 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. ${ }^{18}$

[^7]
### 1.5.1 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.

### 1.5.2 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 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 1.5.5.

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

[^8]

Figure 1.2: Overall Design
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 1.2 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. ${ }^{21}$

### 1.5.3 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

[^9]

Figure 1.3: 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.
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' answers, and were monitored by surveyors in the room as depicted in Figure 1.3 and Figure 1.4. 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.

[^10]

Figure 1.4: 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.

### 1.5.4 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 1.2, 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 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.

Table 1.2: Balance Tests for women and men at baseline

| Control variable | Treatment <br> Mean | Control <br> 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. ${ }^{*} \mathrm{SC} / \mathrm{ST} / \mathrm{OBC}$ represent castes in India. Standard errors are clustered at the class level. p-value for joint test of significance was 0.38 .

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 A. 1 and tables A.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 pvalue 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 A. 5 and at the class level in table A.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 A. 7 and table A. 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 A.4. Next, I discuss how data was collected for various outcomes in the surveys.

### 1.5.5 Data

## 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 1.5.1, 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.1.1. 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 in this paper. I also undertook placebo exercises to show that stigma is not driving the results on sexual harassment reported in section 1.6.1. 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 1.4 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. ${ }^{26}$
[^11]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))..$^{2728}$ These measures are also more comprehensive than those used by Demographic and Health Surveys to collect data on intimate partner violence for women.

## 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.

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. (2019a); 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

[^12]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. ${ }^{31}$

## 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 harassment. 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.1.2.

[^13]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.1.4 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.1.3

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 signups 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.

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.1.5. 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

[^14]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.1.4. 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.

### 1.5.6 Econometric specification

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

$$
\begin{equation*}
Y_{i c g}=\beta_{1} T_{c g}+\beta_{2}^{\prime} X_{i c g}+\beta_{3}^{\prime} K_{c g}+\alpha_{g}+\gamma_{s}+\epsilon_{i c g} \tag{1.1}
\end{equation*}
$$

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, $T_{c g}$ is whether the class c in college g was assigned to receive the male intervention or not, $X_{i c g}$ are student characteristics, $K_{c g}$ are class characteristics taken from administrative data, $\alpha_{g}$ are college level fixed effects, $\gamma_{s}$ are strata (sex ratio $\times$ field of study $\times$ year of study) fixed effects following standard practise (Glennerster and Takavarasha (2013)). $Y_{i c g}$ 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. $\beta_{1}$ 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:

$$
\begin{gather*}
Y_{m c g}=\beta_{1} T_{c g}+\beta_{2} \text { Mixed_Gender }{ }_{m c g}+\beta_{3} T_{c g} \times \text { Mixed_Gender }_{m c g}+\beta_{4}^{\prime} X_{m c g}  \tag{1.2}\\
+\beta_{5}^{\prime} K_{c g}+\alpha_{g}+\gamma_{s}+\epsilon_{m c g}
\end{gather*}
$$

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 $\left(Y_{\text {mcg }}=1\right)$ rather than opting to do it alone $\left(Y_{m c g}=0\right)$. MixedGender ${ }_{m c g}$ 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. $X_{m c g}$ are student level controls and $K_{c g}$ 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.

$$
\begin{gather*}
Y_{i c}=\beta_{1} \text { Female_treatment }_{i c}+\beta_{2} \text { High_Intensity }_{c}+  \tag{1.3}\\
\beta_{3} \text { High_Intensity }_{c} \times \text { Female_treatment }_{i c}+\gamma_{s}+\beta_{4} X_{i c}+\epsilon_{i c}
\end{gather*}
$$

$Y_{i c}$ 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 $H_{i g h}$ Intensity $_{c}$ is a dummy that takes value 1 if class $c$ was assigned to the high intensity treatment. $\beta_{3}$ 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, $\beta_{2}$ 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 $\beta_{1}$ gives the effect of being treated in a low intensity class as compared to someone who is untreated in a low intensity class. $\gamma_{s}$ are strata fixed effects. Standard errors are clustered at the class level and controls are selected by post-double selection LASSO method.

### 1.6 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.

### 1.6.1 Results for the male intervention

## 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 1.2. I look at the effects of the training on these three different types of events and then look at an overall index -same class indexthat combines all types together. The results are shown in table 1.3. 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 A. 10 and table A.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 A.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 1.5. With the caveat that these studies focus only on IPV and other crimes, my results for sexual harassment are close to Green et al. (2020) and Abramsky et al. (2014) who find negative effects of 0.06 and 0.03 , respectively. ${ }^{35}$

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. ${ }^{36}$

[^15]Table 1.3: Women's self-reported exposure to sexual harassment perpetrated by men in their class

| Sexual Harassment | Same Class <br> Index <br> $(1)$ | Mild <br> events <br> $(2)$ | Intermediate <br> events <br> $(3)$ | Extreme <br> events <br> $(4)$ |
| :--- | :---: | :---: | :---: | :---: |
| 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 |  |  |  |  |
| RI p values | - | $[0.315]$ | $[0.204]$ | $[0.009]^{* * *}$ |
| N | $[0.061]^{*}$ | $[0.482]$ | $[0.263]$ | $[0.007]^{* * *}$ |
| Control mean (Non-standardized) | 1255 | 1195 | 1165 | 1165 |

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. ${ }^{*} \mathrm{p}<0.1$, ${ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$. Asterisks denote significance: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.


Figure 1.5: 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.

## 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 1.4, 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 1.5.5. This is what I find in Figure 1.6 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 1.4, 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

[^16]Table 1.4: Effects of the male intervention on opposite sex relationships

|  |  | Survey Measures |  | Lab-in-Field |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | Opposite sex <br> relationship <br> index <br> $(1)$ | Dating <br> - | Opposite <br> sex friends | Switch away <br> (same sex) | Stick to <br> (opp sex) |
|  | $(2)$ | $(3)$ | $(4)$ | $(5)$ |  |
| 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 |  |  |  |  |  |
| Control Mean | 1895 | 1539 | 1810 | 838 | 531 |
| All women | 0.00 | 0.02 | 0.11 | 0.40 | 0.61 |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.
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 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 A.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 A.13. Lastly, I show in table A. 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.

## Mechanism outcomes

In this section, I study the main mechanisms highlighted in section 3.1 to explain the effects I find on sexual harassment and opposite sex relationships through the lens of the theoretical framework.

In table 1.5, 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

[^17]

Figure 1.6: 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.
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, 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 1.7, 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 1.6 for men. However, there was a significant decline in women's perception of legal costs to sexual harassment by

[^18]

Figure 1.7: Treatment effects on awareness of men X-axis lists the different types of events on which I tested men's awareness. Yaxis reports the percentage of men who answered the question correctly. Red bars represent the treatment classes and blue bars are the control classes.

Table 1.5: Effect of male sexual harassment awareness training on men and women

|  | Overall <br> $(1)$ | Sit 1 <br> $(2)$ | Sit 2 <br> $(3)$ | Sit 3 <br> $(4)$ | Sit 4 <br> $(5)$ | ICC <br> $(6)$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| 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 |  |  |  |  |  | 1423 |
| Control Mean | 0.00 | 1624 | 1580 | 1904 | 1904 |  |
| All women |  | 0.93 | 0.86 | 0.39 | 0.74 | 0.14 |
| 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. Sit. is abbreviation for situation. 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: ${ }^{*} \mathrm{p}<0.1,{ }^{*}{ }_{\mathrm{p}}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.
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 A. 15 and table A.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 A.17, table A.20, and table A.21. I find that for each of the situations, men report higher 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.

Table 1.6: Mechanism outcomes

|  | Perceived social <br> disapproval <br> index <br>  <br>  | Perceived <br> legal costs <br> index <br> $(1)$ | Victim Blaming <br> Attitudes <br> Attitudes | Proportion <br> sign-up <br> for volunteership |
| :--- | :---: | :---: | :---: | :---: |
| All men |  |  |  | $(4)$ |

Note: Regression results from estimating equations for dependent variables in columns on classlevel 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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 1.6. 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 1.5.5, 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 1.6. 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.

### 1.6.2 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.

Table 1.7: Effects of female intervention on sexual harassment

| Sexual | Sexual harassment <br> from same class <br> index | Mild <br> events | Intermediate <br> events | Extreme <br> events |
| :--- | :---: | :---: | :---: | :---: |
|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| Female Treatment $\beta_{1}$ | 0.0728 | 0.0247 | -0.0009 | 0.0144 |
|  | $(0.0994)$ | $(0.0334)$ | $(0.0251)$ | $(0.0229)$ |
| High intensity $\beta_{2}$ | 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 $\beta_{3}$ | $(0.1169)$ | $(0.0421)$ | $(0.0395)$ | $(0.0272)$ |
|  |  |  |  |  |
| 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 1.4. Clustered standard errors are in parenthesis and strata fixed effects are included in all columns. PDSLASSO is used for selecting controls. Asterisks denote significance: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.001$.

## Impact on sexual harassment

I show results from estimating the regression in equation 1.3 in table 1.7. First, the estimate for $\beta_{3}$ 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.

## Opposite sex relationships

Next, I show results from a similar specification as above for opposite sex relationships index in table 1.8. 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 $4 \mathrm{p} . \mathrm{p}$. 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.

## Mechanism outcomes

Finally, I look at mechanisms for both men and women together in a coefficient plot in figure 1.8. 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.

### 1.7 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 be-

Table 1.8: Effects of female intervention on opposite gender relationships

|  |  | Survey Measures |  |  | Lab in Field |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Opposite | Dating | Opposite |  | Switch away | Stick to |
| gender | (same class) | sex |  | same | opposite |  |
| index | - |  | friends |  | sex | sex |
|  | $(\mathbf{1})$ | $(2)$ | $(3)$ |  | $(4)$ | $(5)$ |
| Female Treat $\beta_{1}$ | -0.0902 | $-0.0440^{* *}$ | -0.0231 |  | -0.0214 | 0.0520 |
|  | $(0.0956)$ | $(0.0171)$ | $(0.0266)$ |  | $(0.0956)$ | $(0.0816)$ |
| High intensity $\beta_{2}$ | 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 Treat $\beta_{3}$ | $(0.1192)$ | $(0.0377)$ | $(0.0435)$ |  | $(0.1090)$ | $(0.1086)$ |
|  |  |  |  |  |  |  |
|  | 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 $\mathbf{1}$ is an index using Anderson method that combines answers to each variable in columns $\mathbf{2}$ to $\mathbf{5}$. Dependent variable is standardization, in $\mathbf{2}$, 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 $\mathbf{4}$, of dummy equal 1 if switch when paired with same gender partner in the game and in $\mathbf{5}$ 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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.


Figure 1.8: 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.
haviour. 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. ${ }^{41}$

[^19]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.
(2014); Sudarshan and Bhattacharya (2009)).

## Bibliography

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011): "Reference points and effort provision," American Economic Review, 101, 470-92.

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.

Alan, S., S. Ertac, E. Kubilay, and G. Loranth (2020): "Understanding gender differences in leadership," The Economic Journal, 130, 263-289.

Alan, S., S. Ertac, and I. Mumcu (2018):"Gender stereotypes in the classroom and effects on achievement," Review of Economics and Statistics, 100, 876-890.

Ambuehl, S. and S. Li (2018): "Belief updating and the demand for information," Games and Economic Behavior, 109, 21-39.

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 $\mathcal{E}^{3}$ 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.

Azmat, G. and N. Iriberri (2010): "The importance of relative performance feedback information: Evidence from a natural experiment using high school students," Journal of Public Economics, 94, 435-452.

Azmat, G. and B. Petrongolo (2014): "Gender and the labor market: What have we learned from field and lab experiments?" Labour Economics, 30, 32-40.

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.

Banerjee, R., N. D. Gupta, and M. C. Villeval (2020): "Feedback spillovers across tasks, self-confidence and competitiveness," Games and Economic Behavior, 123, 127-170.

Barber, B. M. and T. Odean (2001): "Boys will be boys: Gender, overconfidence, and common stock investment," The quarterly journal of economics, 116, 261-292.

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.

Benjamin, D. (2018): "Errors in Probabilistic Reasoning and Judgment Biases," Working Paper.

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.

Blanes i Vidal, J. and M. Nossol (2011):"Tournaments without prizes: Evidence from personnel records," Management Science, 57, 1721-1736.

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.

Block, R. A. and D. R. Harper (1991): "Overconfidence in estimation: Testing the anchoring-and-adjustment hypothesis," Organizational Behavior and Human Decision Processes, 49, 188-207.

Bohnet, I., A. Van Geen, and M. Bazerman (2016): "When performance trumps gender bias: Joint vs. separate evaluation," Management Science, 62, 1225-1234.

Bohren, J. A., A. Imas, and M. Rosenberg (2019): "The dynamics of discrimination: Theory and evidence," American Economic Review, 109, 3395-3436.

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 (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 D. Y. Yang (2021):"Misperceptions about Others,".
Buser, T., M. Niederle, and H. Oosterbeek (2014): "Gender, competitiveness, and career choices," The quarterly journal of economics, 129, 1409-1447.

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 28233338.

Carlana, M. (2019): "Implicit stereotypes: Evidence from teachers' gender bias," The Quarterly Journal of Economics, 134, 1163-1224.

Castagnetti, A. and R. Schmacker (2020): "Protecting the Ego: Motivated Information Selection," Working Paper.

Charness, G. and U. Gneezy (2012): "Strong evidence for gender differences in risk taking," Journal of Economic Behavior E Organization, 83, 50-58.

Charness, G., R. Oprea, and S. Yuksel (2018): "How Do People Choose Between Biased Information Sources? Evidence from a Laboratory Experiment," Working Paper.

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

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree - An open-source platform for laboratory, online, and field experiments," Journal of Behavioral and Experimental Finance, 9, 88-97.

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 (2019a): "Stereotypes and Belief Updating," Working Paper.
_ (2019b): "When to Apply?" Working Paper, Harvard Business School.
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.

Croson, R. and U. Gneezy (2009): "Gender differences in preferences," Journal of Economic Literature, 47, 448-74.

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,"

Danz, D., A. J. Wilson, and L. Vesterlund (2020): "Belief Elicitation: Limiting Truth Telling with Information or Incentives," CESifo Working Paper.

Dean, J. T. and S. Jayachandran (2019): "Changing Family Attitudes to Promote Female Employment," in AEA Papers and Proceedings, vol. 109, 13842.

Deaux, K. and T. Emswiller (1974):"Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female." Journal of Personality and Social Psychology, 29, 80-85.

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.

Dobrescu, L. I., M. Faravelli, R. Megalokonomou, and A. Motta (2021): "Relative performance feedback in education: Evidence from a randomised controlled trial," The Economic Journal, 131, 3145-3181.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," Journal of the European Economic Association, 9, 522550.

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.

Dweck, C. S., W. Davidson, S. Nelson, and B. Enna (1978): "Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis." Developmental Psychology, 14, 268-276.

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

Eil, D. and J. M. Rao (2011): "The good news-bad news effect: asymmetric processing of objective information about yourself," American Economic Journal: Microeconomics, 3, 114-38.

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.

Espinoza, P., A. B. A. da Luz Fontes, and C. J. Arms-Chavez (2014): "Attributional gender bias: Teachers' ability and effort explanations for students' math performance," Social Psychology of Education, 17, 105-126.

Exley, C. L. and J. B. Kessler (2018):"Motivated Errors," Tech. rep.
— (2019): "The gender gap in self-promotion," Tech. rep., National Bureau of Economic Research.

Falk, A., F. Kosse, H. Schildberg-Hörisch, and F. Zimmermann (2020): "Self-assessment: The role of the social environment," Working Paper.

Falk, A. and F. Zimmermann (2016): "Beliefs and utility: Experimental evidence on preferences for information," IZA Discussion Paper No. 10172.

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.

Flory, J. A., A. Leibbrandt, and J. A. List (2015): "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions," The Review of Economic Studies, 82, 122-155.

Folke, O. and J. K. Rickne (2020): "Sexual Harassment and Gender Inequality in the Labor Market," CEPR Discussion Paper No. DP1473\%.

Ganguly, A. and J. Tasoff (2016): "Fantasy and dread: the demand for information and the consumption utility of the future," Management Science, 63, 4037-4060.

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

Gneezy, U. and J. Potters (1997): "An experiment on risk taking and evaluation periods," The Quarterly Journal of Economics, 112, 631-645.

Goldin, C. (2014): "A grand gender convergence: Its last chapter," American Economic Review, 104, 1091-1119.

Golman, R., D. Hagmann, and G. Loewenstein (2017): "Information avoidance," Journal of Economic Literature, 55, 96-135.

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.

Gurdal, M. Y., J. B. Miller, and A. Rustichini (2013): "Why blame?" Journal of Political Economy, 121, 1205-1247.

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. $2562 \%$.

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 SelfEmployed 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," .

Jones, E. E. and V. A. Harris (1967):"The attribution of attitudes," Journal of Experimental Social Psychology, 3, 1-24.

Jones, T. M., K. V. Fanson, R. Lanfear, M. R. Symonds, and M. Higgie (2014): "Gender differences in conference presentations: A consequence of selfselection?" PeerJ, 2, e627.

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

Karlsson, N., G. Loewenstein, and D. Seppi (2009): "The ostrich effect: Selective attention to information," Journal of Risk and Uncertainty, 38, 95-115.

Karni, E. (2009): "A mechanism for eliciting probabilities," Econometrica, 77, 603-606.

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 ViolenceEvidence from a Field Experiment in Ethiopia," .

Landsman, R. (2019): "Topics in Labor and Experimental Economics," Ph.D. thesis, University of Pittsburgh.

Lindo, J. M., P. Siminski, and I. D. Swensen (2018): "College Party Culture and Sexual Assault," American Economic Journal: Applied Economics, 10, 23665.

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.

Masatlioglu, Y., A. Y. Orhun, and C. Raymond (2017):"Intrinsic information preferences and skewness," Working Paper.

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," .

Michelmore, K. and S. Sassler (2016): "Explaining the gender wage gap in STEM: Does field sex composition matter?" RSF: The Russell Sage Foundation Journal of the Social Sciences, 2, 194-215.

Milkman, K., M. Akinola, and D. Chugh (2013): "Discrimination in the Academy: A Field Experiment," SSRN, Working Paper.

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

Mobius, M., M. Niederle, P. Niehaus, and T. Rosenblat (2014): "Managing self-confidence: Theory and experimental evidence," Working Paper.

Montanari, G. and S. Nunnari (2019): "Audi Alteram Partem: An Experiment on Selective Exposure to Information," Working Paper.

Moore, D. A. and P. J. Healy (2008): "The trouble with overconfidence," Psychological Review, 115, 502.

Moore, D. A., S. A. Swift, Z. S. Sharek, and F. Gino (2010): "Correspondence bias in performance evaluation: Why grade inflation works," Personality and Social Psychology Bulletin, 36, 843-852.

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

Niederle, M. and L. Vesterlund (2007): "Do women shy away from competition? Do men compete too much?" The Quarterly Journal of Economics, 122, 1067-1101.

Nielsen, K. (2018): "Preferences for the resolution of uncertainty and the timing of information," Working Paper.

Oster, E., I. Shoulson, and E. Dorsey (2013): "Limited life expectancy, human capital and health investments," American Economic Review, 103, 19772002.

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.

Ross, L. (1977): "The intuitive psychologist and his shortcomings: Distortions in the attribution process," in Advances in experimental social psychology, Elsevier, vol. 10, 173-220.

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.

Sarsons, H. (2017): "Recognition for group work: Gender differences in academia," American Economic Review, 107, 141-45.

- (2019): "Interpreting signals in the labor market: evidence from medical referrals," Working Paper.

Selody, K. (2010): "Board independence and the gender pay gap for top executives," Working Paper.

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2015): "Financial attention," The Review of Financial Studies, 29, 863-897.

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.

Zimmermann, F. (2014): "Clumped or piecewise? Evidence on preferences for information," Management Science, 61, 740-753.

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

## 2 Demand for Information by Gender: Evidence a Laboratory Experiment

We study preferences for seeking information about one's own ability within a maletyped task. Using a lab experiment, we study how women and men might differ in their preferences for information and how it is affected by the features of the information environment. We consider a setting where first men and women solve a male-typed task, and then choose to receive information about their performance from less or more informative feedback structures. We then vary the costs and benefits of seeking more informative feedback which was also incentivized. In a between-subjects experiment, we make the informative feedback i) private to the subject, ii) public, iii) subjectively judgemental and iv) strategic. We find that women are less likely to opt for more informative feedback than men by about 20 p.p. in the first three treatments but this gap vanishes in the strategic treatment. Interestingly, however, this is driven entirely by men who change their behaviour due to strategic incentives. Our results have implications for how men and women learn differently in male-typed domains.

### 2.1 Introduction

Information about one's own past performance can be important to have well calibrated beliefs, induce higher effort, increase productivity, social learning and selfconfidence (Dobrescu et al., 2021; Azmat and Iriberri, 2010; Blanes i Vidal and Nossol, 2011; Banerjee et al., 2020; Falk et al., 2020). This can have repercussions for outcomes like wages. In many environments like workplaces or academic environments, information in the form of feedback can be given to students or employees via different information environments. In this paper, we are motivated by information environments where feedback can be noisy and may or may not come with greater public visibility, subjectivity, or strategic incentives. For instance, in an academic setting a graduate student can receive feedback through presentations that allows him or her to receive more informative feedback. But the public nature of presentations can also provide them with opportunities to advertise their skills to new collaborators in the audience or receive positive views from their colleagues. This can, however, also have negative consequences if the presentation does not go well or the student signals low ability due to noise in the feedback process or shocks like technical glitches, or poor preparation. This can incentivize one to opt for less informative feedback only from their supervisor (or manager in a workplace) which is less public. What the student or worker thinks about his or her own ability might affect their choice of feedback in such cases.

In this paper, we are motivated to study how men and women might differ in their demand for information within different feedback environments. Women and men might behave differently in terms of information seeking, for instance, in maletyped domains. Women are documented to hold low beliefs about their ability or performance in male stereotypical fields, like maths and science (Bordalo et al., 2019; Coffman, 2014). If women start with low beliefs about their ability in male-typed domains then they may opt to receive less informative feedback if they low confidence or choose more informative feedback if they recognize the benefits that it can have for their growth. Other features of the feedback environment like those mentioned above can further affect demand for feedback. Understanding how women and men might differ in their demand for feedback would be crucial to understand whether or how they form beliefs about their own ability. These beliefs can have adverse impact on gender gaps in self-promotion (Exley and Kessler, 2019), contribution of ideas (Coffman, 2014) and job applications (Coffman et al., 2019b). This is crucial especially in light of gender wage gap that has persisted even after controlling for occupation and human capital differences between men and women (Goldin, 2014). In this paper, we use a laboratory experiment in which subjects perform a maletyped task and get the choice to receive more or less information about their own
performance. We focus on a male-typed task because male-dominated industries and sectors tend to be more remunerative and competitive (Azmat and Petrongolo, 2014).

In our experiment, all participants perform a male-stereotypical task and are ranked according to their performance relative to all other subjects in the same experimental session. They are then asked to state their prior beliefs of being in the top-half. Subjects then choose to receive information, in the form of signals, from one out of two information structures: one being more informative than the other. They are informed that their choice will be payoff-relevant since they would be asked to report their posterior beliefs, which were elicited in an incentive-compatible fashion. Our experiment features a between-subject design in which we vary the costs and benefits of receiving information from the most informative information structure. In this way, we can study how information preferences are shaped by the environment in which it is sought. In the first treatment, the private feedback treatment (T1), both information structures disclose information privately. This means that only the subject gets to receive feedback about her own performance. In this sense, there is only an incentive to seek more information. This helps to first understand whether there are initial differences in information seeking preferences.

Then, motivated by common economic and social settings on how individuals might access information, we vary the costs and benefits of receiving information from the most informative informational structure. We designed a public feedback only treatment (T2) where another person in the session observes the feedback that the subject received; while in the public feedback with judgment treatment (T3) this other person also sends a judgment to the subject about her performance. These treatments help to understand whether the publicness of performance feedback and/or the possibility of receiving a written judgment from another person affects men's and women's information seeking behavior by changing the (psychological) costs connected to demanding information. For instance, in an academic setting, Jones et al. (2014) show that women are less likely to want to present their work in academic conferences and less likely to present for higher amount of time. These settings are crucial to receive feedback about one's work and, thus, the results suggest that the publicness of feedback may drive (part of) these differences by gender. Finally, in the public feedback with hiring decision treatment (T4), the other person has the option to hire the subject if the latter chooses the more informative public feedback. Hiring the subject is optimal if she is in the top half, whereas the subject profits if she is hired regardless of her relative ranking. Thus, T4 introduces incentives related to seeking public information. T4 is modelled after environments where the publicness of information comes with the opportunity to be
hired for future projects or opportunities. ${ }^{1}$
Since we incentivized subjects to receive information, we expect that a subject should demand for information unless they think they have a severe aversion to receiving that information. This aversion could be if they have lower beliefs and confidence in their ability and/or if they want to protect their ego (Castagnetti and Schmacker, 2020). A person who thinks they are of high ability will be more likely to opt for more feedback if it is made public, if they expect it to come with subjectively good judgement or if they think that it can get them better opportunities (strategic incentives). These incentives are reversed if the subjects have a low belief about themselves or of their own performance.

We find first that there are significant differences in the prior beliefs of men and women. Men on average believe that they are $61 \%$ likely to be in the top half of ability distribution while for women it is $46 \%$. We then see stark gender differences in information seeking behavior across treatments. First, a significant fraction of women choose the less informative information structure in T1. Around $20 \%$ of women choose the less informative information structure, whereas only $3.6 \%$ of men do so. Notice, that subjects avoid information even though more information is optimal as subjects know that accuracy in posterior beliefs is incentivized. This finding strongly suggests that only women are avoiding performance feedback. Shares of male and female subjects choosing the less informative information source are not statistically different in T2 and T3 compared to those in T1 although our coefficients are bit large. The magnitudes, however, are insufficient to remove the gender gap as confirmed by our statistical tests. ${ }^{2}$ We find that T4 increases the overall share of subjects who choose the less informative structure. Importantly, this is driven mainly by men thereby removing the gender gap in information seeking. Share of men choosing the less informative information structure increases to $21.9 \%$ in T4 which is six times the share in T1. This suggests that men (and not women) avoid performance feedback strategically to increase their chances of being hired.

In this paper, we study whether men and women differ in their demand for feedback about their performance within a male-typed task or domain. Gender wage gaps in male-typed sectors or domains are especially large. For instance, (Michelmore and Sassler, 2016) show that gender wage gaps are higher in more male typed sectors within STEM like engineering. Male dominated industries also tend to be more remunerative and competitive (Azmat and Petrongolo, 2014). (Exley and Kessler, 2018) also focus on a male-typed task in their study due to similar reasons.

[^20]Large perceived gender differences in ability in male-typed domains can exacerbate selection into such domains. For instance, men perceive themselves to be of a higher ability in finance and also dominate the financial industry (Barber and Odean, 2001). Differences in men and women's performance in maths and science has been the focus of attention in many studies like (Niederle and Vesterlund, 2007) and Buser et al. (2014). Even after holding ability constant, studies have found that the gender gap in beliefs in male-typed domains can have substantial consequences for academic, financial or career track choices (Buser et al. (2014); Barber and Odean (2001); Flory et al. (2015)). Our results inform discussion about how men and women might differ in how much they learn about their own ability in a male dominated environment and how that might affect further their beliefs about themselves.

Our paper contributes to several streams of literature. First, we contribute to the literature on information preferences and, more specifically, on information avoidance. ${ }^{3}$ Our T1 experiment builds on Castagnetti and Schmacker (2020) who study information choices between structures that vary in informativeness. T2, T3, and T4 deviate from them to understand role of costs and benefits of receiving information to study their implications for demand for information. Further, our task is male stereotypical especially allowing us to study the influence of male stereotypical nature of the task on information preferences. ${ }^{4}$ Relatedly, we contribute to the literature on information structure selection where we focus on information related to one's own performance. This literature studies individual preferences for information that either has instrumental value (Charness et al., 2018; Montanari and Nunnari, 2019; Ambuehl and Li, 2018) or not (Falk and Zimmermann, 2016; Masatlioglu et al., 2017; Nielsen, 2018; Zimmermann, 2014).

Different papers have shown that individuals have a preference to avoid information that may carry negative news even if the information is costless and can aid in making more informed decisions. In the context of performance feedback, experimental studies have found that a significant fraction of individuals avoid information about their relative rank (Eil and Rao, 2011; Mobius et al., 2014). We contribute to this literature in several ways. First, we study how these preferences are shaped by different costs and benefits connected to demanding (more) information. Thus, we look at how the publicness of the information and an opportunity of being "hired" shape the demand for performance feedback. Next, our experimental task is purposely chosen to be male stereotypical as opposed to the above-mentioned literature. For instance, (Eil and Rao, 2011) provide information about two variables: the phys-

[^21]ical attractiveness of the subjects (as rated by the opposite sex), and their ranking of IQ in the distribution of these variables within a session while (Mobius et al., 2014) provide information about subjects performance in an IQ task. In our experiment, the subjects performed a task that included questions on assembling objects, sports, maths, general sciences, and mechanical comprehension which are perceived to be male stereotypical, in particular, maths, science and sports (Bordalo et al. (2019)). Lastly, we test whether men and women have differential preferences in seeking information (or avoiding it) within a male stereotypical domain. We believe that this will directly speak to various policy settings or debates which try to understand how women perform in male stereotypical settings and what may or may not hinder their progress. ${ }^{5}$

Second, we contribute to a large, relatively recent and thriving literature on gender stereotypes. This literature emphasizes the role of stereotypes on both selfassessments and the evaluation of others (Alan et al., 2018; Bohnet et al., 2016; Bordalo et al., 2019; Carlana, 2019; Coffman, 2014; Coffman et al., 2019a; Milkman et al., 2013). It shows that male stereotypical tasks lead individuals to believe that women's performance is below average. Information about own performance may be crucial in such cases for accurate self-perception. We add to this literature by studying how men and women might learn differently about themselves when they have to select into receiving more or less information about their performance. If people hold biased beliefs about their own performance in stereotypical tasks, demanding information might be crucial to correct these beliefs. However, we find that women are less likely to self-select into receiving more information about their own performance reducing the accuracy of their beliefs. This is different from uniquely studying how individuals react to information that is exogenously provided to them as in (Coffman et al., 2019a). Importantly, we also shut down any channel of anticipated gender discrimination by others since no information about subjects' gender is revealed in public feedback treatment. ${ }^{6}$ Relatedly, we add to the literature on gender differences in preferences by providing novel evidence of gender differences in preferences for information. ${ }^{7}$ This literature has shown that men and women differ in their preferences for risk (e.g., see Charness and Gneezy (2012)), leadership (e.g., Alan et al. (2020)), competition (Niederle and Vesterlund, 2007), among other dimensions.

The rest of the paper is organized as follows. In Section 2, we describe our experimental design. In Section 3, we report the results of our experiment. In

[^22]Section 4, we discuss our results and in section 5 we conclude.

### 2.2 Experimental Design

Key features of the experiment are as follows. First, we asked subjects to perform a test. Second, we asked them to choose between two different information structures from which to receive information about their own relative performance. The only difference between the information structures is how much information they reveal. Third, we varied the implications of choosing the most informative information structure in a between-subject design. That is, whether it provides the information: a) privately; b) publicly (to another person in the session); c) publicly with possibility of receiving good versus bad judgment from another person; d) publicly with possibility that another person makes a hiring decision based on subject's performance, which monetarily affects the subject. ${ }^{8}$ As explained in section 2.2.6, subjects signed up for different sessions and the sessions were then assigned to different treatments. Thus the treatment assigned to different subjects is as good as random. In our results later, we also control for baseline controls that were collected to minimize any differences between sessions that occur by chance.

### 2.2.1 The Quiz and Prior Beliefs

At the outset of the experiment, participants were asked to solve a quiz. It consisted of 25 questions. The questions were equally split into one of these five categories: assembling objects, general science, maths, mechanical comprehension, and sports. Subjects were given 10 minutes to answer, in any order, as many questions as they could. We incentivized subjects to exert effort in the quiz. We paid them £2.00 per correct answer out of three randomly chosen questions. Since any three questions could be randomly selected, it helped to ensure against concerns that subjects answered only a certain number of questions to reach a target level of earnings. On average, the subjects answered 13 questions correctly, and the distribution by treatment is provided in figure B.17.

We then elicited subjects' beliefs about the likelihood with which they thought that their score in the quiz fell in the top half among all participants in the session. ${ }^{9}$ This belief elicitation stage was incentive compatible. We paid subjects either $£ 6.00$ or nothing based on the accuracy of their answers following the probabilities matching mechanism (Karni, 2009). We chose this method because it does not rely on risk preferences nor on subjective expected utility. ${ }^{10}$ While we explained to subjects

[^23]the main implications of this method (i.e., that it was in their best interest to report truthfully what they really believed), we did not explain the mechanics of the procedure. We did this as withholding the description of the mechanism increases truthful reporting (see Danz et al. (2020)). ${ }^{11}$

### 2.2.2 Choices of Information Structures

Participants were then randomly matched with one other subject in their session. Subsequently, they were asked to choose between two different information structures for receiving information (feedback) about their performance. Irrespective of the choice, performance feedback came in the form of noisy but informative signals. That is, if the subject was in the top (bottom) half of the distribution, she was more likely to correctly receive information about her being in the top (bottom) half. But one of the information structure was more accurate than the other.

Figure 2.1 shows the two information structures. Once the subject chose one information structure, a signal (a ball), corresponding to the subject's relative position in the test, was drawn from the urn. The subject was then shown the signal drawn (a green or a red ball). A green ball with a ' + ' sign meant that the subject was in the top half of the distribution and a red ball with a '-' sign meant they were in the bottom half of the performance distribution. As it can be seen, the information structure on the right is more informative than the one on the left. This is because the precision of information structure B is higher than A's. If the subject was in the top half then information structure B provided them with a green ball $90 \%$ of the time and information structure A showed green ball only $60 \%$ of the time. If the subject was in the bottom half then probabilities were the same but now for a red ball instead of green.

At the time subjects made this choice, they were informed that they would be asked their (posterior) beliefs about their relative ranking again and would be paid according to how accurately they predict it. This meant that acquiring information about their performance was payoff-relevant as it could increase the accuracy of their reported posterior beliefs. However, we varied the implications of choosing information structure B in a between-subjects design. In particular, our experiment features four conditions: (T1) private feedback, (T2) public feedback without judgment, (T3) public feedback with judgment and (T4) public feedback with a hiring choice. We now explain them in turn.

[^24]Figure 2.1: Information Structures in the Experiment


## (T1) Private feedback

In this treatment, the signal received, irrespective of the information structure chosen, was privately seen by the individual. Hence, if subjects choose the less informative information structure (A) in this treatment, this can be attributed to a preference for avoiding performance feedback.

## (T2 \& T3) Public feedback without and with judgement

In the public feedback without judgment, choosing information structure B implied that the signal received by the subject is also revealed to one of the subject's session member. ${ }^{12}$ In the public feedback with judgment, on top of the publicness of the signal from information structure B, the matched partner was asked to send the subject a written message. In particular, the partner had to choose between two predetermined messages: 1) "Your performance must have been really good relative to others", and 2) "Your performance must have been awful relative to others". These messages were deliberately personal to add more subjectivity than the baseline feedback.

In these treatment variations we aim to capture individuals' (un-)willingness to demand performance feedback when it implies that others have access to it too. Moreover, we can also learn whether these preferences are shaped when others could make subjective remarks about the performance.

[^25]
## (T4) Public feedback with hiring decision

In this treatment, the signal from information structure $B$ was not only revealed to the partner but also allowed the partner to make a hiring decision. In particular, after seeing the subject's signal, her partner could choose whether to "hire" her or not. If the subject was hired, then the subject earned $£ 6.00$ and nothing otherwise. Importantly, the partner had an incentive to hire a subject who would be in the top half. If he hired a participant who was in the top (bottom) half of the distribution, he would earn $£ 6.00$ ( $£ 0.00$ ). If instead the partner chose not to hire her he would then earn $£ 2.50$ for sure. On the other hand, if the subject chose the information structure A, then the partner was not asked to make a hiring decision and both players earned $£ 2.50$ in this part of the experiment. This mimics the real world where not showing one's performance to others might reduce hiring opportunities. ${ }^{13}$ This treatment, thus, added a monetary implication of choosing the most informative information structure. The screenshot of how this information was communicated is provided in appendix figure B.14, figure B. 15 and figure B.16.

This treatment allows us to shed light on another force that might prevent people from demanding performance feedback. That is, individuals might have a preference for avoiding performance feedback if that increases their expected returns. In other words, we shed new light on whether strategic considerations affect the demand for information.

In Table 2.1 we provide a summary of the treatment variations and we indicate the main features of each treatment.

Table 2.1: Main Features of the Treatments

| Treatment | Public Signal | Judgment | Hiring Decision |
| :--- | :---: | :---: | :---: |
| (T1) Private Feedback |  |  |  |
| (T2) Public Feedback without Judgment | $\checkmark$ |  |  |
| (T3) Public Feedback with Judgment | $\checkmark$ | $\checkmark$ |  |
| (T4) Public Feedback with Hiring Decision | $\checkmark$ |  | $\checkmark$ |

Subjects were asked to complete a comprehension questionnaire consisting of five questions to ensure that they understood the main features of the information structures and what they entailed for informativeness of the feedback, the publicness of the signals and their partner's set of actions. They could not proceed to the information structure selection until they answered these questions correctly.

[^26]
### 2.2.3 Signal Received and Posterior Beliefs

Subjects then received the signal from their chosen information structure. They were then asked to state their belief about being in the top $50 \%$. We again incentivized subjects to report their beliefs truthfully with the same belief elicitation procedure explained in Section 2.2.1.

Please note that since the two members in the pair were facing the same decisions, it meant that in the public feedback treatments, each subject was shown the signal of her partner if she/he chose information structure B. In the public feedback with judgment she was asked to send the partner one of the two predetermined messages as well; while in the public feedback with hiring decision, she was asked to decide whether to hire the partner or not as well. Across all treatments, subjects were first shown the signal for their own performance as per their chosen information structure and asked about their posterior beliefs (which was incentivized). They were then shown the signal of their partner (if the partner chose the public feedback) and asked to send the judgement statement (in T3) or to make the hiring decision (in T4). Then finally they were shown the judgement statement chosen by their partner (in T3) or the hiring decision made by him/her (in T4) if the subject had chosen information structure B.

Importantly, to avoid hedging motives across the different payoff-relevant parts of the experiment (including the quiz payments), at the end of the experiment only one these parts was randomly selected to actual count for payments. We informed participants about this feature of the payment scheme at the outset of the experiment.

### 2.2.4 Signal Received and Posterior Beliefs

Subjects then received the signal from their chosen information structure. They were then asked to state their belief about being in the top half in their session. We again incentivized subjects to report their beliefs truthfully with the same belief elicitation procedure explained in section ??

Please note that since the two members in the pair were facing the same decisions, it meant that in the public feedback treatments, each subject was shown the signal of her partner if she/he chose information structure B. In the public feedback with judgment she was asked to send the partner one of the two predetermined messages as well; while in the public feedback with hiring decision, she was asked to decide whether to hire the partner or not as well.

Across all treatments, subjects were first shown the signal for their own performance as per their chosen information structure and asked about their posterior
beliefs (which was incentivized) as shown in figure ??. They were then shown the signal of their partner (if the partner chose the public feedback) and asked to send the judgement statement (in T3) or to make the hiring decision (in T4). Then finally they were shown the judgement statement chosen by their partner (in T3) or the hiring decision made by him/her (in T4) if the subject had chosen information structure B.

To avoid hedging motives across the different payoff-relevant parts of the experiment (including the quiz payments), we randomly selected one part of the experiment to count for final payments. This means that the participants were paid for either the quiz, the prior belief elicitation or the posterior belief elicitation. We informed participants about this feature of the payment scheme at the outset of the experiment. This helps to alleviate the concern that participants may exert effort in only one or two parts of the experiment once they think they have performed well or reached a certain financial goal within the experiment.

### 2.2.5 Debriefing

Participants were then asked to answer some debriefing questions. First, we asked subjects to explain in their own words why they chose one information structure over the other. We paid them $£ 0.50$ for their answers. Next, we elicited, in an incentive compatible fashion, participants' risk preferences with the Gneezy and Potters (1997) risk elicitation task. We then asked subjects to complete a questionnaire that measures individuals' overconfidence in terms of over precision. ${ }^{14}$ It is based on Block and Harper (1991) and it asks' participants to specify a confidence interval to their answers such that there is a $90 \%$ chance that each answer falls inside it. Finally, we asked them to answer a questionnaire. It included demographic questions such as participants' age, country of origin, and gender. We also asked them whether they had participated in experiments before and if they knew anyone in the session. Participants were finally asked a general willingness to take risk question (Dohmen et al., 2011).

### 2.2.6 Implementation

The experiment was conducted in fall 2019 at the Economics laboratory of Warwick University. Overall, 344 participants, recruited through the SONA recruitment software, took part in the experiment. We conducted 24 sessions of about 45 minutes each. Participants earned an average payment of $£ 11.00$, including the show-up of

[^27]£5.00. We programmed the experiment with oTree (Chen et al., 2016). Descriptive statistics of the sample are provided in Appendix B.2.

### 2.3 Results

We first discuss the summary statistics of our sample and then look at the prior beliefs. We then look at the information structure choice by treatment and gender. In this analysis, we also perform econometric regressions to quantify our results and show the robustness of our findings. Finally, we look at posterior beliefs to study how subjects learnt from the signals and formed their posterior beliefs. We then discuss the implications for efficiency in this context.

### 2.3.1 Descriptive Statistics

The summary of the characteristics of our sample is provided in appendix table B.2. The summary statistics show that between $61 \%$ to $52 \%$ of the experimental subjects were women for a mean of $56 \%$. The average age of the subjects was 21 since these were university students, majority of the students had a quantitative education (depending on which department you belonged to), and most of the subjects are in the middle of the risk preference elicitation range. That is, most perceive themselves as not that willing to take risk. ${ }^{15}$ and subjects were willing to bet close to $50 \%$ of the amount in a risky lottery. On average, the subjects think that there is a $53 \%$ chance that they will be in top half of the performance distribution, close to $50 \%$ had a previous experience with doing lab experiments and on average $29 \%$ knew others in the lab. These characteristics are balanced across treatments for all variables except for the share who knew others in the lab. Close to $60 \%$ knew each other in sessions that were assigned to treatment 1 while it was close to $22 \%$ in the other sessions. In any randomized experiment there can be imbalances by chance, but to make sure that these differences are not driving our results, we show results when controlling for all these demographic variables together. The variables we include in the regressions are : age, education (year of study in the university), department you are enrolled in, the mother tongue, country of origin, ethnicity, mother tongue, previous experience with experiments, share of other subjects they know in the experiment, their own perception of risk aversion, risk choice, and their prior belief of their own performance. We show the questions we asked for demographics in the experiment in appendix section B.2.1.

[^28]
### 2.3.2 Prior Beliefs

To measure prior beliefs, we asked subjects their belief about being in the top $50 \%$ of the performance distribution in the session. The mean prior belief across all individuals is 53.157 (s.d. 22.236). However, as expected due to the gender stereotype of the task, prior beliefs for men are higher than those for women. While for men the average belief is 61.669 (s.d. 20.460), it is only 46.497 (s.d. 21.324) for women. We control for prior beliefs given these gender differences. In this way, we can also assess whether gender effects in information structure choices are driven by prior beliefs.

Importantly, we can also look at how accurate prior beliefs are. To do this, we can study the distance between the prior belief and participants' true rank. For those in the top half, this means the distance between $100 \%$ and their prior belief, whereas for those that are in the bottom half, the distance is exactly their prior belief. That is, for subjects in the top half the (lack of) accuracy is [(I (In the top half)=1)-Prior Belief] and for those in the bottom half the error is [(I(In the top half)=0)-Prior Belief] where prior belief is the probability with which participant thinks that he/she is in the top half of the distribution in their session. The average distance is 42.826 (s.d. 21.280) and it is essentially the same across genders: 42.397 (s.d. 22.305) for men and 43.161 (s.d. 20.495) for women.

If subjects were fully calibrated in their beliefs, the average distance should be zero; whereas if they knew nothing it should be around $50 \%$. We see that the mean distance is somewhere in the middle but substantially closer to $50 \%$. In sum, what this analysis strongly suggests is that there is scope for learning and the subsequent choice on information structures is relevant.

### 2.3.3 Information Structure Choice

We start by providing an overview of information structure choices by treatment and gender. Table 2.2 shows the proportion of subjects who chose information structure A over B. Interestingly, men and women react very differently to the treatment conditions. For men we can see that as we move from (T1) to (T4) the proportion of subjects choosing A over B increases from $3.6 \%$ to $21.9 \%$. While for women the proportion of those who choose information structure A is around $20.0 \%$ in all treatments.

We now perform the regression analysis to quantify our experimental results. We start by looking at information structure choices in (T1) and by gender. In particular, we start with the following specification:

Table 2.2: Gender differences in the choice of information structure A by treatment.

|  | Proportion <br> Men | Proportion <br> Women | All |
| :--- | :---: | :---: | :---: |
| (T1) Private Feedback | 0.036 | 0.236 | 0.152 |
| (T2) Public Feedback w/o Judgment | 0.061 | 0.169 | 0.128 |
| (T3) Public Feedback w. Judgment | 0.102 | 0.188 | 0.147 |
| (T4) Public Feedback w. Hiring Decision | 0.219 | 0.224 | 0.222 |
| All Treatments | 0.108 | 0.195 | 0.163 |

The table provides the proportion of participants who choose information structure A by treatment and gender.

$$
\begin{equation*}
Y_{i}=\alpha+\beta_{1} \text { Female }_{i}+X_{i}^{\prime} \beta_{2}+\epsilon_{i} \tag{2.1}
\end{equation*}
$$

$Y_{i}$ is a dummy variable which takes the value 1 if participant i chooses information structure A. Female $e_{i}$ is a dummy taking the value 1 if participant i is a female. $X_{i}$ refers to control variables for individual i. Here, $\beta_{1}$ is the difference in take up of the less informative information structure between men and women in (T1). This coefficient informs us on gender differences in take up of the less informative information structure. Robust standard errors are reported in the parenthesis.

Table 2.3 shows the results. In Columns (1)-(3) we report the results of OLS regressions, while in columns (4)-(6) we report those for probit regressions. ${ }^{16}$ In Columns (2) and (5) we add subjects' prior beliefs as control variable, while in Columns (3) and (6) we further add the demographic control variables. We report marginal effects for probit specifications for ease of interpretation.

We can see that women are 20 percentage points more likely to choose the less informative information structure (A). The coefficients are stable across all specifications after including the prior beliefs and the rest of the controls mentioned previously. In sum, in (T1) there is a highly significant gender difference in take-up of information. We can also test whether the proportion of individuals that chose information structure A over B for each gender is equal to zero. While for men we find that it is not statistically different significant from zero ( p -value=0.326), for women it is statistically different from zero ( p -value $=0.013$ ). This shows that, despite more informative signals being monetarily incentivized, women have a preference for avoiding information about their own performance in the quiz.

In appendix table B.3, we show results with session fixed effects. Although there were only 24 sessions, and the fixed effects necessarily reduce the degrees of freedom

[^29]especially for treatment 1 which has a total of 66 subjects, we find that our results remain significant (although now are marginally significant). The coefficients are slightly lower than before but indicate that women are close to 18 p.p more likely to demand less informative feedback than men. We also clustered the standard errors at the session level and find that the results stay qualitatively similar to the above for both linear and the probit models although for OLS we lose the power to detect the effects when we add all the controls. These results are in appendix table B.6.

Table 2.3: Choice of Feedback Mode B in (T1)

|  | OLS |  |  |  | Probit |  |  |
| :--- | :--- | :---: | :--- | :--- | :--- | :--- | :--- |
|  | $(1)$ <br> Info B | $(2)$ <br> Info B | $(3)$ <br> Info B |  | $(4)$ <br> Info B | $(5)$ <br> Info B | $(6)$ <br> Info B |
|  |  |  |  |  |  |  |  |
| Female | $0.201^{* *}$ | $0.220^{* *}$ | $0.238^{* *}$ | $0.201^{* * *}$ | $0.210^{* * *}$ | $0.175^{* *}$ |  |
|  | $(0.079)$ | $(0.085)$ | $(0.115)$ | $(0.079)$ | $(0.079)$ | $(0.073)$ |  |
|  |  |  |  |  |  |  |  |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |  |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |  |
| R-Squared | 0.077 | 0.098 | 0.205 | 0.105 | 0.131 | 0.313 |  |
| N | 66 | 66 | 66 | 66 | 66 | 66 |  |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue, country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.

Now, to understand the impact of the publicness of the feedback (signals) on information seeking behavior, we run the following specification:

$$
\begin{equation*}
Y_{i j}=\alpha+\beta_{1} \text { Female }_{i j}+\beta_{2} T 2_{j}+\beta_{3} T 3_{j}+\beta_{4} T 2_{j} \times \text { Female }_{i j}+\beta_{5} T 3_{j} \times \text { Female }_{i j}+\beta_{6} X_{i j}+\epsilon_{i j} \tag{2.2}
\end{equation*}
$$

We augment Equation (2.1) to include variables $T 2$ and $T 3$, which are dummy variables taking the value of 1 if the treatment was (T2) or (T3), respectively. We also include their interaction with the female dummy ( $\mathrm{T} 1 \times$ Female and $\mathrm{T} 2 \times$ Female). $\beta_{2}\left(\beta_{3}\right)$ captures differences in take up of the less informative information structure between T 2 (T3) and T1 for men. $\beta_{4}\left(\beta_{5}\right)$, on the other hand, captures the difference in the impact of introducing the publicness of the signal in T2 (and judgement in T3) in information structure B on take-up of less informative feedback between women and men.

Table 2.4 shows the results. As before, in Columns (1)-(3) we report the results

Table 2.4: Choice of Information Structure B in (T1)-(T3)

|  | OLS |  |  | Probit |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | (1) <br> Info B | (2) <br> Info B | $\begin{gathered} (3) \\ \text { Info B } \end{gathered}$ | (4) <br> Info B | (5) <br> Info B | (6) <br> Info B |
| Female | $\begin{aligned} & 0.201^{* *} \\ & (0.078) \end{aligned}$ | $\begin{aligned} & 0.197^{* *} \\ & (0.081) \end{aligned}$ | $\begin{aligned} & 0.197^{* *} \\ & (0.084) \end{aligned}$ | $\begin{aligned} & 0.213^{* *} \\ & (0.090) \end{aligned}$ | $\begin{aligned} & 0.210^{* *} \\ & (0.091) \end{aligned}$ | $\begin{aligned} & 0.203^{* *} \\ & (0.081) \end{aligned}$ |
| T2 | $\begin{aligned} & 0.025 \\ & (0.055) \end{aligned}$ | $\begin{aligned} & 0.026 \\ & (0.055) \end{aligned}$ | $\begin{aligned} & 0.066 \\ & (0.061) \end{aligned}$ | $\begin{aligned} & 0.055 \\ & (0.129) \end{aligned}$ | $\begin{aligned} & 0.057 \\ & (0.129) \end{aligned}$ | $\begin{aligned} & 0.114 \\ & (0.132) \end{aligned}$ |
| T3 | $\begin{aligned} & 0.066 \\ & (0.056) \end{aligned}$ | $\begin{aligned} & 0.067 \\ & (0.056) \end{aligned}$ | $\begin{aligned} & 0.092 \\ & (0.062) \end{aligned}$ | $\begin{aligned} & 0.118 \\ & (0.117) \end{aligned}$ | $\begin{aligned} & 0.121 \\ & (0.117) \end{aligned}$ | $\begin{aligned} & 0.149 \\ & (0.111) \end{aligned}$ |
| T2 $\times$ Female | $\begin{aligned} & -0.092 \\ & (0.103) \end{aligned}$ | $\begin{aligned} & -0.097 \\ & (0.102) \end{aligned}$ | $\begin{aligned} & -0.126 \\ & (0.105) \end{aligned}$ | $\begin{aligned} & -0.087 \\ & (0.094) \end{aligned}$ | $\begin{aligned} & -0.092 \\ & (0.090) \end{aligned}$ | $\begin{aligned} & -0.097 \\ & (0.062) \end{aligned}$ |
| T3 $\times$ Female | $\begin{aligned} & -0.114 \\ & (0.105) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.112 \\ & (0.103) \end{aligned}$ | $\begin{aligned} & -0.115 \\ & (0.073) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.071) \end{aligned}$ | $\begin{aligned} & -0.096 \\ & (0.054) \end{aligned}$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |
| p value $\beta_{1}+\beta_{4}=0$ | 0.02 | 0.02 | 0.09 | 0.03 | 0.04 | 0.06 |
| p value $\beta_{1}+\beta_{5}=0$ | 0.17 | 0.24 | 0.11 | 0.19 | 0.28 | 0.06 |
| R -Squared | 0.036 | 0.037 | 0.104 | 0.049 | 0.049 | 0.160 |
| N | 254 | 254 | 254 | 254 | 254 | 254 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.
of OLS regressions, while in columns (4)-(6) we report the marginal effects at the mean for probit regressions. In Columns (2) and (5) we add subjects' prior beliefs as control variable, while in Columns (3) and (6) we add the demographic control variables. From columns (1), (2) and (3) we can see that the female dummy coefficient remains stable and statistically significant at the $5 \%$ level regardless of the specification and the addition of control variables.

We do not find that there is an additionally significant impact of treatments (T2) and (T3) on information structure selection between men and women. We find that women stay between 10.9 to 8 p.p more likely to opt for less informative feedback. We have also added the f tests in the table. This is especially true for T 2 and for T3 in some of the specifications. Results from probit regressions, reported
in columns (4)-(6), are similar. The coefficients on the interactions are also not statistically significant. These results suggest that making a signal public (with or without judgement) does not change significantly the preference of women for less informative feedback.

The results with session fixed effects are in appendix table B. 4 and with clustered standard errors are in appendix table B.7. With session fixed effects, our coefficients go down slightly and become marginally significant but our results stay similar qualitatively. This is intuitive since session fixed effects affect our degrees of freedom and reduce the power. With standard clustering our results stay the same, but with wild cluster bootstrap we find that our OLS estimates either become marginally significant or lose it but stay below significance of $15 \%$. On the other hand our probit estimates stay significant much like the session fixed effects. Overall, we conclude that our results indicate that T 2 and T 3 have low impact on the gender gap in demand for information we established in T1.

Finally, we study the impact of adding strategic considerations in (T4) on information preferences. Before running any regression, in Figure 2.2 we provide graphical evidence of information structure choice by gender. In the figure, we pool together the results from treatments (T1), (T2), and (T3). ${ }^{17}$ From the figure, we can see that gender differences in the choice of information structures are stark between men and women in the non strategic treatments while in the strategic environment (T4) the gender gap fully disappears. The difference shrinks from 0.122 ( p -value $=0.006$ ) to 0.005 ( p -value $=0.956$ ).

[^30]Figure 2.2: Take Up of Information Structure B by Strategic/Non Strategic Environment.


Notes: Gender differences in take up of private feedback in the strategic and not strategic treatments.

To quantitatively analyze the impact of the strategic treatment on information structure choices, we now run the following regression:

$$
\begin{align*}
Y_{i j}=\alpha+ & \beta_{1} \text { Female }_{i j}+\beta_{2} T 2_{j}+\beta_{3} T 3_{j}+\beta_{4} T 4_{j}+\beta_{5} T 2 \times \text { Female }_{i j}  \tag{2.3}\\
& +\beta_{6} T 3 \times \text { Female }_{i j}+\beta_{7} T 4 \times \text { Female }_{i j}+X_{i j}^{\prime} \beta_{8}+\epsilon_{i j}
\end{align*}
$$

Similar to Equation 2.2, we look at the additional impact that the strategic environment in (T4) has on the choice of the less informative information structure compared to (T1) for men. $\beta_{4}$ provides this information. Moreover, $\beta_{7}$ captures how and whether the strategic incentives (of being hired) affect the gender gap in information seeking behavior.

The results are shown in Table 2.5. ${ }^{18}$ We find that (T4) leads to a large increase in take up of private feedback for men compared to T1 (the omitted category). This increase amounts to approximately 18 to 19 percentage points in columns 1, 2 and 3. While, as we saw before, there is no differential impact of (T2) or (T3) on gender gap in take-up of private feedback, here we find a statistically significant decrease in women's take-up of private feedback due to T4 as compared to men. The coefficients against $T 4 \times$ Female are almost equal to the coefficient against Female ( 20 percent-

[^31]age points), thereby, entirely removing the gender gap. An F-test for whether the sum is zero of the coefficients against the female dummy and the interaction with (T4) cannot be rejected. As we saw in Table 2.2 and Figure 2.2, this result is entirely driven by men (also apparent from the estimates for the $\beta_{4}$ coefficient, which gives the effect of strategic considerations on men's take up of private information). This shows that the strategic incentive treatment increases take-up of private (less informative) feedback only for men indicating men's sensitivity to (strategic) incentives. Probit estimates show a similar story. ${ }^{19}$ The coefficients against (T4) are highly significant for probit specifications. Women, on the other hand, continue to demand less informative feedback and do not respond to such incentives.

We provide results with session fixed effects and clustered standard errors in appendix tables B. 5 and tables B.8. With session fixed effects our coefficient on gender stays significant although marginal and the value becomes smaller. We still cannot reject that the gender gap is different from zero as given by the F-test at the bottom which is our key finding in the main results above. However the fixed effects reduce the power for us to detect significance on the strategic treatment for men while the coefficients stay large. With clustering we find, that most of our coefficients stay marginally significant. Crucially we cannot reject the null that overall the gender gap gets removed under the strategic interactions treatment using wild cluster bootstrap on the sum of coefficients $\left(\beta_{1}+\beta_{7}\right)$.

We also show the differences in any of the social demographics of those who chose to get feedback from different information structures in appendix table B.2. It shows that apart from gender, the two groups were similar along all other dimensions. $69 \%$ of those who chose information structure B were women, while it only $53 \%$ for those who chose information structure A. This confirms that other than gender, the other controls are not driving our results despite there being imbalance for one of our controls at the baseline.

### 2.4 Discussion

### 2.4.1 Posterior Beliefs

We now study individuals' posterior beliefs. The mean posterior belief is 53.360 (s.d. 29.935). While it is 55.894 (s.d. 31.676) for men, it is 51.378 (s.d. 28.425) for women.

We now look at the accuracy of participants' posterior beliefs. That is, whether beliefs are closer to $100 \%$ ( $0 \%$ ) for participants in the top (bottom) half. We look at this distance depending on the information structure chosen. The average distance

[^32]Table 2.5: Overall difference in take up of less informative feedback, all treatments

|  | OLS |  |  | Probit |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | (1) <br> Info B | $\begin{gathered} (2) \\ \text { Info B } \end{gathered}$ | (3) <br> Info B | $\begin{gathered} (4) \\ \text { Info B } \end{gathered}$ | $\begin{gathered} (5) \\ \text { Info B } \end{gathered}$ | (6) <br> Info B |
| Female | $\begin{aligned} & 0.201^{* *} \\ & (0.078) \end{aligned}$ | $\begin{aligned} & 0.198^{* *} \\ & (0.080) \end{aligned}$ | $\begin{aligned} & 0.2011^{* *} \\ & (0.083) \end{aligned}$ | $\begin{aligned} & 0.237^{* *} \\ & (0.100) \end{aligned}$ | $\begin{aligned} & 0.235^{* *} \\ & (0.101) \end{aligned}$ | $\begin{aligned} & 0.242^{* *} \\ & (0.095) \end{aligned}$ |
| T2 | $\begin{aligned} & 0.025 \\ & (0.055) \end{aligned}$ | $\begin{aligned} & 0.026 \\ & (0.055) \end{aligned}$ | $\begin{aligned} & 0.040 \\ & (0.059) \end{aligned}$ | $\begin{aligned} & 0.062 \\ & (0.148) \end{aligned}$ | $\begin{aligned} & 0.064 \\ & (0.149) \end{aligned}$ | $\begin{aligned} & 0.088 \\ & (0.150) \end{aligned}$ |
| T3 | $\begin{aligned} & 0.066 \\ & (0.056) \end{aligned}$ | $\begin{aligned} & 0.067 \\ & (0.056) \end{aligned}$ | $\begin{aligned} & 0.082 \\ & (0.060) \end{aligned}$ | $\begin{aligned} & 0.138 \\ & (0.142) \end{aligned}$ | $\begin{aligned} & 0.139 \\ & (0.142) \end{aligned}$ | $\begin{aligned} & 0.175 \\ & (0.141) \end{aligned}$ |
| T4 | $\begin{aligned} & 0.184^{* *} \\ & (0.074) \end{aligned}$ | $\begin{aligned} & 0.186^{* *} \\ & (0.075) \end{aligned}$ | $\begin{aligned} & 0.198^{* * *} \\ & (0.075) \end{aligned}$ | $\begin{aligned} & 0.294^{*} \\ & (0.159) \end{aligned}$ | $\begin{aligned} & 0.299^{*} \\ & (0.159) \end{aligned}$ | $\begin{aligned} & 0.356^{* *} \\ & (0.157) \end{aligned}$ |
| T2 $\times$ Female | $\begin{aligned} & -0.092 \\ & (0.103) \end{aligned}$ | $\begin{aligned} & -0.096 \\ & (0.102) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.094 \\ & (0.099) \end{aligned}$ | $\begin{aligned} & -0.098 \\ & (0.096) \end{aligned}$ | $\begin{aligned} & -0.111 \\ & (0.075) \end{aligned}$ |
| T3 $\times$ Female | $\begin{aligned} & -0.114 \\ & (0.105) \end{aligned}$ | $\begin{aligned} & -0.117 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.110 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.124^{*} \\ & (0.075) \end{aligned}$ | $\begin{aligned} & -0.127^{*} \\ & (0.073) \end{aligned}$ | $\begin{aligned} & -0.126^{* *} \\ & (0.058) \end{aligned}$ |
| T4 $\times$ Female | $\begin{aligned} & -0.196^{*} \\ & (0.119) \end{aligned}$ | $\begin{aligned} & -0.202^{*} \\ & (0.118) \end{aligned}$ | $\begin{aligned} & -0.227^{*} \\ & (0.121) \end{aligned}$ | $\begin{aligned} & -0.162^{* * *} \\ & (0.051) \end{aligned}$ | $\begin{aligned} & -0.164^{* * *} \\ & (0.049) \end{aligned}$ | $\begin{aligned} & -0.173^{* * *} \\ & (0.094) \end{aligned}$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |
| P -value $\beta_{1}+\beta_{7}=0$ | 0.955 | 0.963 | 0.775 | 0.954 | 0.938 | 0.979 |
| R-Squared | 0.033 | 0.034 | 0.0375 | 0.042 | 0.043 | 0.103 |
| N | 344 | 344 | 344 | 344 | 344 | 344 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue, country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.
is now 38.942 (s.d. 28.015); ${ }^{20}$ it is 38.087 for those who chose the most informative information structure and it is 43.339 (s.d. 25.201) for those who chose the less informative one. Comparing this distance to that found in their priors, we see that the distance has significantly reduced for those subjects in the first group ( $\Delta=4.664$, p-value $=0.005$ ), while it has not for those in the latter group $(\Delta=1.750$, p -value $=0.503$ ).

[^33]This analysis shows that participants that chose information structure A do not learn as much compared to those that chose information structure B. Thus, subjects who chose the most informative information structure ended up with more accurate posterior beliefs. This highlights that the way that individuals select into receiving information has strong implications in terms of how much they learn and, subsequently, on how accurate their (posterior) beliefs are. Gender differences in take-up of less versus more informative feedback, thus, implies gender differences in accuracy of these beliefs about own ability.

### 2.4.2 Implications for Efficiency

In terms of efficiency, we can see that in treatments (T1) to (T3) the rational payoff maximizing choice is to receive information from information structure B. In fact, choosing information structure A is inefficient as it provides less information and, as we have seen, it leads to more inaccurate posterior beliefs. We find that women are more likely to choose the inefficient choice avoiding information on their true rank.

Recall in T 4 it was crucial whether one is revealed to be above or below median to one's partner (if information structure B is chosen) for payoffs. Therefore, to analyse the efficiency implications of choosing one information structure over the other, we divide our sample of men and women into two groups: above median and below median performers. In particular, we look at above median performers for whom choosing more informative information structure is the efficient choice both in terms of learning one's true rank and of being hired by the partner. ${ }^{21}$ We find that above median performers seek lesser information under strategic than non strategic environments implying that inefficiency rises in this context. Quantitatively, take up of more information goes down for above median performers by 11 p.p. in the strategic environment (T4) compared to the non strategic treatments (T1-T3). As discussed above, however, strategic incentives do remove the gender gap between men and women but only by increasing men's take up of lesser information. Thus, a reduction in gender gap comes at the cost of increased inefficiency overall, and specifically for men.

### 2.4.3 Overall results

In a between-subjects design, we asked experimental subjects to perform a male stereotypical task and then choose to receive feedback about their performance from more or less informative information structures. We find that women are less likely

[^34]to choose to receive more informative feedback about their own performance than men. This gender gap is of about 20 p.p. and remains stable across most of our specifications. Women's demand for less informative feedback stays at approximately $20 \%$ whether we make more informative feedback public, subjective or strategic indicating a sticky and low demand for information. We cannot rule out null effects of introducing publicness to the feedback or subjective judgement to the more informative feedback, although our coefficients are high relative to the pure gender gap. But our tests show that the these two treatments are not able to change the gender gap in demand for information, in particular when feedback becomes public. Our coefficients for subjective feedback are much larger and we cannot rule out its effects on gender gaps for all specifications. When we introduce strategic incentives, we detect significant effects of the strategic incentives treatment in closing the entire gender gap in demand for information. However, this is driven by an increase in percentage of men who demand less informative feedback from $3 \%$ to $11 \%$ between T1 to T4.

The literature on gender stereotypes has shown that stereotypes can affect selfassessments and self-confidence in gender incongruent domains and drive gender differences in behaviour. We had elicited the prior beliefs of our experimental subjects which we control for in our specifications but they do not affect our results at all. On the other hand, when we asked the subjects in the debriefing of why they chose the private feedback, most women answered that they thought that they were in the bottom half and they did not want to others to see it. Thus, although in our regressions, priors do not change the results, the estimate was always negative. That is a higher prior was correlated with less likelihood of less informative feedback.

An experiment with female stereotypical task will allow researchers to understand whether the pattern holds only for a male stereotypical domain or not. To conjecture what might happen, we use qualitative answers we collected from our participants in the debriefing. We show in figure B. 18 that most that both men and women who chose less informative feedback did so because they thought they were of low ability and did not want others to know (in case of T2, T3 or T4). But only women reported that they were not confident in their ability indicating that prior beliefs might play some role. Although, we do not find a role of prior beliefs in our results, it is possible that beliefs might still play a role if there were big differences in prior beliefs. Exley and Kessler (2018) and Bordalo et al. (2019) did not find gender differences in selfperceptions of ability in verbal ability task to model female stereotypical domain. Hence, if modelled through such a task the results we have are likely to hold also within a female-typed domain. Coffman (2014), however, looks at gender differences in tasks that involved quiz questions about art \& literature and entertainment \&
pop culture which are conjectured as more female-typed tasks. The prior differences in beliefs of women's advantage over men in these was lower than that of men over women in male-typed domain captured through sports. ${ }^{22}$ We, thus, hypothesize that the results may be similar to our study since differences in prior beliefs induced by female-typed tasks are not as strong as male-typed tasks if beliefs play a role at all. We leave that for future research.

### 2.5 Concluding Remarks

In this paper, we study whether there are gender gaps in information seeking about own performance in a male stereotypical domain. We find that women are nearly 20 p.p. more likely to choose less informative feedback despite more information being incentivized. We do not find strong evidence that it is affected by adding publicness or judgement to the feedback. However, the gender gap disappears in the treatment where we introduce a strategic consideration of being hired by a partner. This effect is driven entirely by men who increase the take-up of the less informative information structure.

This research provides a new perspective to the literature on gender differences in economic behavior. In particular, it shows that there are gender differences in preferences for information, and that they are shaped by the strategic environment in which it is sought. This, in turn, affects how much someone learns about past performance. This is not without consequences. In fact, this may affect future performance levels through different channels. For example, getting information about your own performance can help you in taking corrective actions for improvement (e.g. higher effort). But it can also help in keeping your ego intact.

Future research should enrich the current analysis at least in two dimensions. First, new research should more closely study the (potentially) many implications of differences in information-seeking behavior by gender. Specially, in male dominated environments in which gender differences in beliefs and outcomes (e.g. wages) are higher. Second, while we provide a careful analysis on how the environment shapes information preferences by gender, we believe that a step forward would be to look at these preferences in real-world settings. Indeed, this may prove particularly fruitful in terms of understanding how the publicness of information shape preferences. In real-world scenarios publicness of the feedback may play a more decisive role compared to what we find here.

[^35]
## 3 Attribution Bias by Gender: Evidence a Laboratory Experiment

In many settings economic outcomes depend on both the competence of the agents involved and on luck. When principals assess agents' performance they can suffer from attribution bias by gender: male agents may be assessed more favorably than female agents because males will be rewarded for good luck, while women punished for bad luck. We conduct a pilot laboratory experiment to study if principals judge agents' outcomes differentially by their gender. Agents perform a task for the principals and performance depends on both the agents' competence and luck. Principals then assess agents' performance and decide what to pay the agents. Our pilot experimental results do not show evidence consistent with attribution bias by gender. While principals' beliefs and payments are heavily influenced by realized outcomes, they do not depend on the gender of the agent. We instead find suggestive evidence that the interaction between the gender of both the principal and the agent plays a role. In particular, principals are more generous to agents of the opposite gender.

### 3.1 Introduction

In most economic settings, outcomes depend on dispositional (internal) factors of the agents involved, such as their effort and ability, as well as on situational (external) factors, such as luck, that are not under their control. This creates room for attribution bias. In psychology, attribution bias is defined as the tendency for people to under-emphasize situational explanations for an individual's observed behavior while over-emphasizing dispositional explanations for their behavior (Ross, 1977). Attribution bias by gender, instead, is understood as the tendency of observers to over-attribute successes to internal factors for males and to external factors to females, while also over-attributing failure to external factors for males and to internal ones for females (Citation Here).

Recent empirical evidence suggests that attribution bias by gender may be at work in several contexts. For instance, Sarsons (2019) shows that, following a patient death, female surgeons are less likely than male surgeons to receive future referrals from other physicians. Moreover, female surgeons experience a smaller increase in future referrals than male surgeons after an unexpectedly successful surgery. Selody (2010) shows similar results but for executive pay in the finance sector. In particular, the study finds that following negative shocks to the firm's performance, female employees' financial rewards are more responsive than rewards to male employees. While, for positive shocks, women's financial rewards are less responsive. Similarly, Landsman (2019) shows that female executives are fired from their jobs at a much higher rate than male executives when there are unexpected industry-wide contractions not specific to the firm. This emerging literature suggests that attribution bias by gender may contribute to these differences in outcomes for men and women. However, many other variables may be at work in these real-world environments that cannot be controlled for - these include agents' real contribution to outcomes, prior experience, and unobserved characteristics, among others. Thus, it is not possible to completely rule out factors other than attribution bias that might drive these differences in outcomes by gender.

In this paper, we present evidence from a pilot laboratory experiment in order to study the presence of attribution bias by gender in a controlled environment. In particular, the goal of the paper is to understand if attribution bias by gender manifests within a principal-agent framework. A lab experiment suits well in understanding the role of attribution bias since it provides a controlled setting in which other factors are unlikely to influence participants' behavior.

Participants in this experiment are first randomly divided into two roles: principals and agents. In each round (out of 20), they are randomly matched into pairs. In each round, agents perform a task for their principals. Principals are rewarded
based on the outcome of the agents' performance, while agents are paid by their principals after the outcome is revealed. Importantly, the outcome produced by the agents is not a deterministic function of their performance, but it also depends on a random component. In each interaction, principals are shown information that allows them to identify the agents' gender. This piece of information is conveyed through the agents' (nick-)names and presented along with other demographic information to minimize demand effects. After each interaction, we elicit agents' and principals' beliefs about the agents' performance.

Our main hypotheses follow from the concept of attribution bias by gender. That is, following a high outcome, we hypothesize that principals will be more likely to attribute it to the agents' internal dispositions if male, while to external factors if female. This will result, in turn, in higher wages being paid to males relative to females conditional on a high outcome. Similarly, we hypothesize that principals will attribute a low outcome to the agents' misfortune if male, while to the agents' dispositions if female. Thus, again we expect that female agents will receive lower payments as compared to male agents conditional on a low outcome.

Our pilot experimental results do not show evidence of attribution bias by gender. While principals' payments are heavily influenced by the realized outcomes, they do not differ by the gender of the agent. Similarly, principals' beliefs about the agents' performance do not differ by gender, although they are heavily influenced by the realized outcomes. In other words, the principals' payments and beliefs are not shaped by the agents' gender. Our pilot results, therefore, suggest that gender is not a driving force when principals assess the agents' performance, at least in a laboratory environment. We find however suggestive evidence that the interaction between principals' and agents' gender affects payment decisions. In particular, principals pay higher wages to agents of the opposite sex.

### 3.2 Literature Review

How individuals attribute causes of behavior and outcomes to both dispositional and situational factors has received considerable attention in the last fifty years in social psychology. In particular, the fundamental attribution error, which refers to the tendency of observers to assign too much weight to dispositional factors (e.g., preferences and ability) while too little to situational factors (e.g., constraints and luck) when interpreting others' behaviours and performances, has been largely studied (e.g., Jones and Harris (1967), Ross (1977), and Moore et al. (2010)). Here, however, we are interested in a specific manifestation of this bias: attribution bias by gender. Espinoza et al. (2014) and Dweck et al. (1978) show within an educational set up that teachers are more likely to attribute success in maths to ability for men and effort for women, while attribute failures or negative evaluations to effort
(non intellectual inadequacies) for men and ability (intellectual inadequacies) for women. Deaux and Emswiller (1974) show same patterns for maths task (male typed task) but do not find the reverse in a female typed task illustrating that type of task may not be important. In this paper, we test whether attribution bias exists within a controlled lab set up and whether it translates into differential payments for unexpectedly bad or good events. We do this by studying beliefs and payments within a principal-agent set-up. We also test the hypothesis using different types of tasks to understand whether prior beliefs could drive attribution bias if at all.

Experimental literature on gender within economics has emphasized gender differences in preferences as explanations of differential economic behaviour and outcomes (labour market, educational choices, career growth etc.) by gender. However, vast literature also shows that discrimination in the environment can also contribute to differences in economic outcomes by gender. This paper is related to several strands of this latter literature. First strand is the literature focusing on gender discrimination as measured in the lab and field. A number of papers have established the presence of gender discrimination in labour markets at the applicant screening stage or at the hiring stage. ${ }^{1}$ Taste based discrimination (when employers derive disutility from hiring women) and statistical discrimination (when gender is used to extrapolate unobservables regarding productivity of potential employees) have been the two most widely studied mechanisms to pin down the drivers of gender differences in labour market outcomes due to discrimination. In an interesting paper, Bohren et al. (2019) show that gender discrimination may be reversed dynamically indicating discrimination driven by biased beliefs. Yet another mechanism that has been studied is that of attribution bias. Sarsons (2019) shows that attribution bias may be one of the channels driving differential impact of unexpected patient deaths (or recoveries) on male and female surgeons' received referrals from a senior physician. The study however cannot pin it down as a source due to challenges in measuring beliefs in observational data. Selody (2010) and Landsman (2019) both show the differential impact of unexpected losses (or gains) on women's labour market outcomes. In particular, these papers provide evidence that women may be punished more due to unexpected losses by way of lower referrals, higher firing rates or lower financial rewards. Sarsons (2017) shows in a different context that women are given lower credit for group work than men and that comes from employers interpreting noisy signals in a biased manner in academia. Using two experiments, the same paper establishes that solo authored papers instead contribute equally to men and women's tenure rates. In solo authored papers however, the paper can only be attributed to the single author. Complementary to these studies, we develop an

[^36]experiment where there is uncertainty about an individual's contribution to output and explicitly model the uncertainty which is known to employers (principals in our set up). We explain in detail in the next section how we use the experiment to pin down attribution bias by gender in the lab.

### 3.3 Experimental Design

An experiment that studies attribution bias by gender requires several ingredients. First, it requires two roles: an agent whose performance is to be evaluated, and a principal who evaluates the agent's performance. Second, the outcome of the agent's performance needs to be a function of both dispositional and situational factors. Third, the principal is aware of the gender of the agent. Our design features all these pieces.

A complete and detailed description of the experiment is presented below. First, we asked participants to fill out a demographic questionnaire. Second, we randomize participants into two roles: principals and agents. In each of the 20 rounds, agents and principals were matched into pairs using the stranger-matching protocol. ${ }^{2}$ Thus, the agent performed a task for the principal. The agent's performance influenced the resulting output, but not deterministically. This output determined the principal's earnings in that round. The principal then proceeded to pay the agent for his performance. In each round, we elicited agents' and principals' beliefs about the agent's contribution to the realized outcome. Finally, subjects were asked to complete a series of questions about the experimental task.

### 3.3.1 The Experiment

At the outset of the experiment, participants were asked to complete a demographic questionnaire that included information about their gender, field of study, level of study, nationality and state of origin, age, caste, and religion. Participants were then randomly assigned into one of two roles: principals and agents. They were informed that these roles were fixed for the whole duration of the session, that the experiment consisted of two tasks, and that the tasks would be played one after the other and for ten rounds each. ${ }^{3}$ We then explained to our participants the structure of a round. While the general features of each round were read aloud by the experimenter, we asked participants to read the specific details on their computer screens. To make sure participants understood the experiment, we encouraged them to ask questions if anything was unclear and we asked them to complete a set of comprehension questions. Participants could not continue with the experiment until

[^37]they had answered all questions correctly.

## Description of a Round

In each round, principals and agents were matched into pairs following the strangermatching protocol. ${ }^{4}$ Participants were informed that, although they could earn money in each of the 20 rounds, at the end of the experiment only one would be randomly selected to count for the final payments.

Agent's performance and output produced At the beginning of each round, the agent performs a task, consisting of a fixed number of questions. These are to be performed in 45 seconds. The agent's performance determines the lottery that is assigned to the principal. Each lottery has only two possible outcomes: High and Low output. The agent's performance (i.e., the number of correctly solved questions) affects the lottery assigned to the principal by increasing the probability that the high output is realized. However, even in the case an agent had solved all questions correctly, there is a positive probability that the resulting output is low.

Principals' payments Once the 45 seconds have passed, the principal is shown the realization of the lottery, which constitutes her payoff for that round (i.e., the output produced by the agent). Importantly, the principal is not informed about the lottery from which the output has been determined and so does not know the number of questions solved correctly by the agent. However, the principal is fully aware of the mapping between the number of correctly answered questions and the lotteries. The principal proceeds by choosing a reward for her agent. In particular, the principal is given access to a pot of $₹ 350^{5}$ and she is free to choose how to divide this amount between her agent, a random agent in the session, and the experimenter. ${ }^{6}$ This separate pot is independent of the realized outcome in that round. Importantly, the agent did not get to see the payment he received until the end of the session. In this way, his (future) performance was not dependent on the history of payments he had received and the principal's payments would not be

[^38]driven by an underlying motive of incentivizing the agent to perform well. Thus, this payment method provides a tighter test for attribution bias by gender.

Principal's beliefs After the payment decision, we elicited the principal's beliefs about the absolute performance of the agent. In particular, we asked the principal to indicate the number of questions that she thought that the agent had solved correctly. We incentivized this question by paying ₹ 50 if the answer was correct. In some sessions we also asked the same question but while the agent was performing the task and so before the outcome of the lottery is realized. That is, we asked the principal to indicate her prior belief about the number of correct questions that the agent would solve correctly. Finally, we also asked the principal two unincentivized questions. First, we asked the principal to guess how many questions she thought that the agent attempted and an hypothetical question on whether she would like to be paired for another round with the same agent.

Agent's beliefs We asked the agent three unincentivized beliefs' questions. First, we asked him to guess the number of questions that he solved correctly. Second, we asked him to guess whether the principal earned the high or low output. Finally, we also asked him his belief about the percentage of the ₹350 that he would receive from the principal.

## Debriefing

Finally, after the two tasks have been completed, we asked participants to answer two sets of questions. First, we asked participants to guess our research questions. Second, we asked participants questions about the previous tasks. For instance, we asked them to assess the difficulty of the tasks and their general feelings during the tasks.

### 3.3.2 Gender Information

In each round, while the agent was performing the task, the principal was shown some demographic information about the agent. In particular, the principal was given information about whether the agent was a university student, his age, and gender. We disclosed gender information of the agent through the means of nicknames. That is, in each round the computer software would assign a gendercongruent nickname to the agent. As the experiment took place in India, the realized nickname was randomly selected from a list of popular Indian names. Since we used only first names, they did not signal caste. All names were the most popular Hindu
names. For instance, female names included "Akansha", "Neha" and "Priya", while for male names these included, among others, "Amit", "Ashish" and "Nitin". ${ }^{7}$

The use of nicknames instead of a direct statement of the agent's gender was implemented to mask the fact that our research question was gender-related and, therefore, to prevent potential distortions due to demand effects and social desirability concerns. Moreover, we opted for nicknames as opposed to real names because we wanted to preserve anonymity and control more carefully for the type of information disclosed via names. For instance, we wanted to make sure that names did reveal the gender of the agent, and that they did not prime religion or caste-related information.

## The Tasks and Output

In each session agents performed two different tasks. However, since we vary these tasks across sessions, we have a total of four tasks: a math task, a Raven task, an effort task, and a memory task. We now provide a description of each of these tasks.

The math task We implemented a variation of the Niederle and Vesterlund (2007) math task. In each round of this task, agents were asked to perform 7 additions. Each addition consisted of three two-digit numbers.

The Raven task In each round of this task, agents were asked to solve three Raven Matrices. In particular for our experiment, we implemented the matrices from the Raven Advanced Progressive Matrices (APM) test. This test is commonly used to measure fluid intelligence.

The effort task For this task we used a variation of the Abeler et al. (2011) effort task. In this task, agents were shown ten $5 \times 5$ matrices that were randomly filled with zeros and ones. Agents were asked to solve as many grids as possible by counting the number of ones in each matrix.

The memory task This task was a working memory exercise. In particular, agents were shown 16 English common words (e.g., cat, umbrella, and house) for 25 seconds. After that, the words disappeared from the screen and they had to write down as many words as they could remember in the following 20 seconds.

[^39]
## Lotteries and Output

As indicated previously, each correct answer in a given task increased the probability of the high output being realized. In each task and for each round we had variation in two dimensions: the mapping of correct answers into the probability of the high output (i.e., the set of lotteries) and the level of the high output. ${ }^{8}$ This variation was randomly determined and orthogonal to each other.

The lotteries Given that the number of questions asked by task differed, the precise mapping of correct questions into the probability of the high output occurring changed by task. However, the overarching feature across tasks was that the probability of the high output was always increasing in the number of correct questions solved by the agent. Moreover, for each task, we had two different mappings: The high and low calibrations. In the former, the probability of the high output started at $50 \%$ had the agent solved one question correctly and, as the agent solved more questions correctly, it could reach and exceed $90 \% .^{9}$ In the latter, instead, the probability of the high outcome started at $5 \%$ and could at most reach $60 \%$ had the agent solved all questions correctly. We varied the mapping to understand whether this feature affects payments, beliefs, and gender-biased attributions.

Output level The high output could take three different levels: ₹ 400 , ₹ 550 , or ₹700. We vary the level of the high output to see whether principals' payments and beliefs are affected by the (potentially) earned high output.

Importantly, both the agents and the principals had access to this information in each round. Agents were shown the mapping and the output level before they performed the task, while principals were shown this information at the time the agents were performing the task. Both agents and principals were given unlimited time to read and process the information, which was provided in table form for intuitive exposition and ease of understanding.

### 3.4 Experimental Hypotheses

Our experiment was designed to analyze whether principals make biased attributions regarding the performance of the agents. In particular, to capture attribution bias we designed an environment in which outputs represent noisy signals of the agent's competence. That is, the output produced in each round is a function of the number of questions answered correctly by the agent (that is, the agent's competence), but also luck. The principal therefore has to base her payment on the basis of the

[^40]lottery's outcome.
In this environment, we test whether the gender of the agent plays a crucial role in the principal's payment and in shaping her beliefs about how much the agent's competence contributed to the output. In particular, our hypotheses follow directly from the concept of attribution bias by gender. We thus conjecture that a principal will attribute a high output to the agent's competence if male, while she will attribute it to luck if female. Similarly, following a low output, the principal will attribute it to misfortune if the agent is male while to competence if the agent is a female. We therefore make the following hypothesis:

The principal's beliefs about the number of correctly solved questions will be higher for male as compared to female agents, following both high and low outputs.

Furthermore, we conjecture that this difference in the principal's beliefs by gender will affect the way principals make their payments to agents. In particular, we hypothesize that:

The principal's payments will be higher for male agents as compared to female agents, following both high and low outputs.

### 3.5 Experimental Results

### 3.5.1 Implementation

The pilot experiment was conducted in July 2018 in the computer lab at Delhi School of Economics. Invited participants belonged to the departments of Economics, Sociology, Commerce and Geography. We recruited a total of 84 subjects and conducted 5 sessions that lasted around 75 minutes each. The participants earned on average ₹ 510 , which includes the show-up fee of ₹ 250 . We programmed the experiment with oTree (Chen et al., 2016).

We begin by examining agents' performance and their beliefs in Section 3.5.2. In particular, we look at agents' performance across tasks and by gender. In this section, we also analyze agents' beliefs about their own performance and their beliefs about their principals' payment decisions. We then investigate principals' payment decisions and their beliefs about their agents' performance in Section 3.5.3. We thus analyse whether principals make biased attributions and payments depending on the gender of their matched agents. In Section 3.5.6 we discuss alternative factors that might be driving our results: the salience of gender information, principals' prior beliefs about the agents' performance, selection of our sample, and whether principals' payments are driven by other demographic information about the agents (i.e. age).

Table 3.1: Mean performance by task and gender

|  | All Agents | Female Agents | Male Agents | P-value Difference |
| :--- | :---: | :---: | :---: | :---: |
| All Tasks | 0.39 | 0.39 | 0.40 | 0.85 |
| Math Task | 0.52 | 0.51 | 0.53 | 0.73 |
| Raven Task | 0.22 | 0.21 | 0.23 | 0.62 |
| Memory Task | 0.33 | 0.33 | 0.33 | 0.85 |
| Effort Task | 0.53 | 0.54 | 0.52 | 0.81 |

Statistical significance is assessed by running regressions of performance on the gender of the agent. Standard errors are clustered at the individual level.

Table 3.2: Mean beliefs about performance by task and gender

|  | All Agents | Female Agents | Male Agents | P-value Difference |
| :--- | :---: | :---: | :---: | :---: |
| All Tasks | 0.51 | 0.49 | 0.52 | 0.41 |
| Math Task | 0.60 | 0.58 | 0.62 | 0.34 |
| Raven Task | 0.47 | 0.46 | 0.48 | 0.78 |
| Memory Task | 0.36 | 0.35 | 0.36 | 0.71 |
| Effort Task | 0.55 | 0.54 | 0.57 | 0.45 |

Statistical significance is assessed by running regressions of performance on the gender of the agent. Standard errors are clustered at the individual level.

### 3.5.2 Agents

## Agents' Performance and their Beliefs

Performance The mean proportion of correctly answered questions across tasks was $39 \%$ (s.d. 0.23). Performance, defined as the proportion of correct questions solved, varies by task: it is highest in the math and effort tasks with over $50 \%$ of questions solved correctly, while lowest for the Raven task with $22 \%$ of correct answers. If we look at performance broken down by gender in Table 3.1, we find no difference in performance by gender across tasks. ${ }^{10}$

Beliefs about performance As can be seen in Table 3.2, agents' beliefs about their own performance differ by task but there are no differences in beliefs by gender. If we compare performance and beliefs, we can see that agents are highly overconfident in the math and the Raven tasks. Indeed, they believe they have solved an excess of around $15 \%$ questions correctly.

[^41]Table 3.3: Mean beliefs about realized outcomes and expected principals' payments

|  | All Agents | Female Agents | Male Agents | P -value |
| :---: | :---: | :---: | :---: | :---: |
| Beliefs about Realized Outcomes |  |  |  |  |
| All Tasks | 0.74 | 0.73 | 0.75 | 0.78 |
| Math Task | 0.77 | 0.76 | 0.77 | 0.87 |
| Raven Task | 0.65 | 0.62 | 0.68 | 0.47 |
| Memory Task | 0.75 | 0.78 | 0.72 | 0.71 |
| Effort Task | 0.87 | 0.85 | 0.88 | 0.72 |
| Beliefs about Principals' Payments |  |  |  |  |
| All Tasks | 0.55 | 0.59 | 0.51 | 0.15 |
| Math Task | 0.59 | 0.60 | 0.58 | 0.74 |
| Raven Task | 0.46 | 0.48 | 0.45 | 0.66 |
| Memory Task | 0.61 | 0.68 | 0.49 | 0.03 |
| Effort Task | 0.0.55 | 0.65 | 0.47 | 0.23 |

Statistical significance is assessed by running regressions of either beliefs about realized outcomes or beliefs about principals' payments on the gender of the agent. Standard errors are clustered at the individual level.

## Agents' Beliefs about Outcomes and Expected Payments

Beliefs about realized outcomes and expected principals' payments At the top of Table 3.3 we look at agents' beliefs about realized outcomes by gender and task. We find essentially the same patterns as compared to beliefs about performance. On the other hand, at the bottom of Table 3.3, we can see that female agents believe that, on average, principals will allocate $59 \%$ of the ₹ 350 to them, while male agents believe it to be around $51 \%$. However, the difference is not statistically significant.

In sum, we find that, while performance differs by task, it does not significantly differ by the agents' gender. Similarly, beliefs about own performance and principals' actions do not differ by the gender of the agent. Male and female agents' beliefs about principals actions differ but not statistically significantly so.

### 3.5.3 Principals

We now turn to study our main outcome variables: The principals' payment decisions and beliefs. We start by considering principals' beliefs and choices depending on the outcome produced by their agents. We then analyze these variables depending on the agents' gender. Importantly, in the following analysis we looked at pooled results that take into account all tasks and calibrations. The outcome produced in each round is coded as 0 when the realized outcome of the lottery was low ( $₹ 0$ ) and 1 if it was high ( $₹ 400$, ₹ 550 , or ₹ 700 ).


Figure 3.1: Principals' payment decisions by realized outcome
Notes: the histograms show the distribution of principals' payment decisions by realized outcome.

## Wages and Principals' Beliefs

Wages In Figure 3.1, we show the distribution of principals' payments depending on the realized outcome. From the figure it is clear that payments depended heavily on the realized outcome: higher payments were made following a high outcome and, vice versa, for low outcomes. ${ }^{11}$ A Mann-Whitney test confirms that the distribution of principals' payments differs significantly by the realized outcome ( p -value $<0.00$ ).

Beliefs Figure 3.2 shows that principals' beliefs ${ }^{12}$ about their agents' performance follow a similar pattern. Principals' beliefs are higher when the output is high as compared to when it is low. A Mann-Whitney test shows the the distribution of principals' beliefs differ significantly by the realized outcome ( p -value $<0.00$ ).

[^42]

Figure 3.2: Principals' beliefs by realized outcome
Notes: the histograms show the distributions of principals' beliefs about the matched agents' performance by realized outcome.


Figure 3.3: Principals' wages by realized outcome and agents' gender Notes: the histograms show the distribution of principals' payment decisions by realized outcome and the gender of the matched agents.

## Wages and Principals' Beliefs by the Gender of their Matched Agents

Wages by gender of the agent We now test our experimental hypotheses put forward in Section 3.4. That is, we analyze whether there are differences in principals' wages depending on the gender of their matched agents. Figure 3.3 shows that, while payments respond to the outcome of the lottery, they do not differentially respond by the agents' gender. Indeed, a Mann-Whitney test fails to reject the null hypotheses of equality in distributions following either a low ( p -value $=0.611$ ) or a high outcome (p-value $=0.883$ ). ${ }^{13}$

Beliefs by gender of the agent The results for beliefs match those for wages. Figure 3.4 shows that, while beliefs are heavily influenced by the realized outcome, they do not shift depending on the agents' gender. Results of a Mann-Whitney test shows no significant different in distributions irrespective of whether the outcome is low or high ( p -value $=0.514$ and p -value $=0.884$, respectively $)$.

[^43]

Figure 3.4: Principals' beliefs by realized outcome and agents' gender Notes: the histograms show the distributions of principals' beliefs about the matched agents' performance by realized outcome and gender of the matched agents.

In sum, our experimental results do not seem to support our experimental hypotheses. In fact, there is no evidence that principals' payments decisions and beliefs are influenced by their agents' gender.

### 3.5.4 Econometric Specification

We next conduct parametric analyses to further analyze the variables affecting the principals' payment decisions. In particular, we estimate the following regression:

$$
\begin{equation*}
Y_{i j}=\beta_{0}+\beta_{1} Z_{i j r}+\beta_{2} \text { Female }_{i}+\beta_{3} \text { Female }_{i} \times Z_{i j r}+\beta_{4} X_{i j r}+\epsilon_{i j} \tag{3.1}
\end{equation*}
$$

where $i$ represents the agent, $j$ is the principal and $r$ is the round, Yij is the dependent variable that is either the principal's payment to the agent or her belief about the agent's performance, $Z_{i j r}$ is the outcome of the lottery produced by agent $i$ matched with principal $j$ in round $r$, Female $_{i}$ is a dummy equal to 1 if the agent is a female and $X_{i j r}$ is a vector of controls and includes principals' demographic variables (age, caste, religious group, field of study, education level, and state) and task characteristics (task, calibration of the lottery, and level of the high outcome). We report standard errors clustered at the principal level in all specifications. $\beta_{2}$ captures whether there are any average differences in payments made to female versus male agents when the outcome produced is low while $\beta_{3}$ captures if there is any difference in payments made to female agent versus male agent in response to a high outcome.

We then also account for the principal's gender to check whether this variable plays any role in the payments made to the agent. We estimate this using the following econometric specification:

$$
\begin{align*}
Y_{i j}= & \beta_{0}+\beta_{1} Z_{i j r}+\beta_{2} \text { Female }_{i}+\beta_{3} \text { Female }_{i} \times Z_{i j r}+  \tag{3.2}\\
& \beta_{4} \text { Female }_{j}+\beta_{5} \text { Female }_{j} \times Z_{i j r}+\beta_{6} X_{i j r}+\epsilon_{i j}
\end{align*}
$$

Female $_{j}$ is a dummy equal to 1 if principal $j$ was a female. $\beta_{4}$ captures if there are any average differences in payments made by male versus female principals in case of a low outcome (holding everything else constant) while $\beta_{5}$ captures whether there is any difference in payments made in response to high versus low outcome by female principal. Our set up allows us to test also for the interaction between agent's and principal's gender causally because of the random matching design of our experiment. Hence we also report estimates from the following specification:

Table 3.4: Regression results for principal's payments

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $101.57^{* * *}$ | $101.59^{* * *}$ | $112.79^{* * *}$ | $110.01^{* * *}$ |
|  | $(14.38)$ | $(17.59)$ | $(33.53)$ | $(33.47)$ |
| Female Agent |  | 1.92 | 1.31 | 10.01 |
|  |  | $(9.25)$ | $(9.06)$ | $(9.54)$ |
| Female Agent $\times$ High Outcome |  | -0.04 | 0.96 | -0.27 |
|  |  | $(13.20)$ | $(12.41)$ | $(12.54)$ |
| Female Principal |  | -34.15 | -33.59 |  |
|  |  | $(40.89)$ | $(39.92)$ |  |
| Female Principal $\times$ High Outcome |  | -15.49 | -12.46 |  |
|  |  |  | $(35.42)$ | $(35.31)$ |
| Same Gender |  |  |  | $-17.22^{* *}$ |
|  |  |  |  | $(6.86)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ |  |
| R-Squared | 0.33 | 0.33 | 0.34 | 0.35 |
| N | 804 | 804 | 804 | 804 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Standard errors are clustered at the principal level.

$$
\begin{array}{r}
Y_{i j}=\beta_{0}+\beta_{1} Z_{i j r}+\beta_{2} \text { Female }_{i}+\beta_{3} \text { Female }_{i} \times Z_{i j r}+\beta_{4} \text { Female }_{j}  \tag{3.3}\\
+ \\
+\beta_{5} \text { Female }_{j} \times Z_{i j r}+\beta_{6} \text { SameGender }_{i j}+\beta_{7} X_{i j r}+\epsilon_{i j}
\end{array}
$$

$\beta_{6}$ here captures whether being matched to an agent of the same gender leads to any differential effect on payments made by principals. The results for principals' payment decisions are shown in Table 3.4.

### 3.5.5 Econometric Results

As it is apparent from the regression Table 3.4 (columns 1 to 4), the outcome of the lottery is the most important and significant variable determining the payment made to the agent. ${ }^{14}$ Going from a low to a high outcome increases the principal's payment to the agent by around ₹ 100 in all the specifications. On the other hand, the agent's gender does not play a role. The coefficient on the female dummy and the interaction with the outcome variable is insignificant and extremely small compared

[^44]Table 3.5: Regression results for principal's beliefs

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.21^{* * *}$ | $0.21^{* * *}$ | $0.16^{* * *}$ | $0.16^{* * *}$ |
|  | $(0.03)$ | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Agent | - | 0.02 | 0.03 | $0.04^{*}$ |
|  | - | $(0.02)$ | $(0.02)$ | $(0.02)$ |
| Female Agent $\times$ Outcome |  | -0.02 | -0.02 | -0.02 |
|  | - | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Principal |  |  | $-0.13^{* *}$ | $-0.13^{* *}$ |
|  | - | - | $(0.06)$ | $(0.06)$ |
| Female Principal $\times$ Outcome |  |  | $0.09^{*}$ | $0.09^{* *}$ |
|  | - | - | $(0.05)$ | $(0.05)$ |
| Same Gender |  |  |  | -0.02 |
|  | - | - | - | $(0.02)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.30 | 0.30 | 0.31 | 0.31 |
| N | 804 | 804 | 804 | 804 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Standard errors are clustered at the principal level.
to the average of the dependent variable which is around ₹ 184 . Thus, principals did not pay differently to women as compared to men conditional on the same outcome. Thus, we cannot reject the null hypothesis that there is no attribution bias by gender.

If we perform the same regressions for beliefs, we find the same patterns. Table 3.5 shows that principals' beliefs are significantly shaped by outcomes while they are not affected by the agents' gender. Female principals are more likely to believe that agents solved a lower proportion of questions correctly in case of a low outcome while they increase their beliefs significantly more than the male principals in response to a high output. However their payments seem to react lower than that of male principals as seen in Table 3.4.

On the other hand, our results show that, while the gender of the agent alone does not influence payments, its interaction with the principal's gender does and significantly so as in column 4 of Table 3.4. In particular, principals payments are significantly higher to agents of the opposite gender, irrespective of the realized outcome. In other words, principals pay around ₹ 17 less to their matched agents if they belong to the same gender. While our data does not allow us to further explore this due to the small sample size, we believe that this result deserves further investigation in future research. In Table 3.5 with beliefs, we find that the coefficient on same gender is negative in line with the evidence for payments, however, this is
not significant.
Taken together, our results show a clear picture where principals' beliefs and payment decisions are heavily and significantly influenced by outcomes. We do not, however, find supporting evidence of gender biased attributions. There is, instead, suggestive evidence that the gender of the agent matters only in interaction with the gender of the principal, for which further mechanisms need to be explored in future research.

### 3.5.6 Robustness Checks

Session and round fixed effects In the appendix we show that our results on principals' payments are robust to including session fixed effects (see Table C.3) and round fixed effects (see Table C.4). Running regressions without controls shows that the main result on principals' payments is robust to including controls (see Table C.5). The results for same gender of the principal and agent also hold indicating a negative effect of being matched with an agent who has the same gender. Similarly, the main results for beliefs are also robust to including session fixed effects (see Table C.8) and round fixed effects (see Table C.9). Results without controls are qualitatively similar and hold in the same direction as shown in Table C.10.

Restricting the sample to the first ten rounds If principals' beliefs in the beginning are much more biased and less likely to have been influenced by looking at past performances, then it is possible that attribution bias by gender was prevalent in the initial rounds. To test for this, we ran similar regressions by restricting sample to the first ten rounds in Table C. 6 and Table C.11. It does not change the interpretation of the results. The results for attribution bias by gender hold as for the rest of the sample. The coefficient on female agent dummy stays small relative to the mean of payments which is ₹ 184 . The coefficient on the interaction of female dummy with high outcome dummy is negative but still small and statistically insignificant. For the case of payments, the coefficient on same gender becomes larger by ₹ 7 than the estimate obtained using the whole sample showing that the effect of being matched to same gender agent may be stronger in the beginning of the experiment. In table C.11, high outcomes no longer make female principals update their beliefs significantly more than males. The results for attribution bias by gender, however, are the same and there is no significant effect of being matched with a female agent on the beliefs of the principal after an unexpectedly high or low outcome. Thus, in summary, if we believe that principals beliefs by gender are stronger in the beginning and may manifest themselves in attribution bias by gender, we find that is not the case.

Removing the first five rounds in the sample We further looked at the results after removing the first five rounds of the sessions to account for the possibility that the participants may not fully understand the experiment in the beginning. in Table C. 7 and Table C.12. We still do not find attribution bias by gender. The coefficients become higher than the main results for female agent but are still statistically insignificant and small in comparison to the mean of payments which is around ₹ 184 . The coefficient for same gender stays high and significant as shown in Table C.7. Thus our results do not seem to be driven by principals not understanding the experiment in the initial few rounds.

## Restricting the sample to different parts of the cumulative distribution

 We estimated equation 1 for different parts of the cumulative distribution for the principals' payments dependent variable. We thus estimate the following econometric specification:$$
\begin{equation*}
Y_{i j x}=\beta_{0}+\beta_{1} Z_{i j r}+\beta_{2} \text { Female }_{i}+\beta_{3} \text { Female }_{j}+\beta_{3} X_{i j r}+\epsilon_{i j} \tag{3.4}
\end{equation*}
$$

The above equation is same as before except that the dependent variable $Y_{i j x}$ is the payment made by the principal when the sample is restricted to include only observations where agents were paid less than or equal to $x$. We vary $x$ from ₹ 50 , ₹ 100 , ₹ 150 , ₹ 200 ₹ 250 and ₹ 300 . The estimates are shown in Table C.13. For all the parts of the cumulative distribution for payments, we find that there is no support for attribution bias by gender in line with our main results. $\beta_{2}$ and $\beta_{3}$ are very small in comparison to the mean payments made in different parts of the distribution mirroring our main findings.

### 3.6 Possible Threats

### 3.6.1 Gender Information

Our lack of experimental evidence for attribution bias by gender could be explained by the way in which we disclosed gender information about the agents. A failure to find attribution bias by gender could be driven by the fact that principals understood that our research question was about gender discrimination and, therefore, they were particularly cautious in preventing such bias from arising during the experiment (e.g., due to a social desirability bias). However, when subjects in the role of principals were asked to guess our research questions at the end of the experiment ${ }^{15}$, no one guessed it was about gender. The most common guesses included answers such as: "the sharing tendency of people", "a study on how individual de-

[^45]cision making is affected when their possible returns are contingent on the actions of another person", and "assessing contracts". ${ }^{16}$

Alternatively, one might worry that displaying gender information using nicknames is not salient enough to induce gender discrimination. While we agree this is a possible interpretation of our results, we believe that the fact that principals did not pay attention to the gender information of the agent is itself a finding. This implies that principals did not judge this piece of information was important in making their payments and attributions. Further, the results regarding the principals' gender in interaction with the gender of the agents shows that principals did pay attention to the agent's gender. In other words, this result provides evidence that information about the agents' gender was salient and principals did take them into account, although in a way not consistent with attribution bias by gender.

### 3.6.2 Prior Beliefs by Gender of the Agent

In two sessions, we also elicited principals' beliefs about the agents' performance prior to any knowledge regarding the realized outcome. While prior beliefs that principals have are slightly higher for male agents than for female agents ( $70 \%$ vs. $67 \%$ ), the difference is not statistically significant ( p -value $=0.29$ ). We therefore do not believe that our results are being driven by differences in prior beliefs.

### 3.6.3 Selection of our Sample

Since we conducted the experimental sessions at Delhi School of Economics, one may wonder whether our "null" results might be driven by selection: women at this university may be positively selected relative to the population. While this is likely to be true, two considerations are worth emphasizing. First, it is not the case that they did better on the tasks than male participants. This means that, at least on our experimental tasks, there is no selection by gender. Second, our sample of highly selected females resembles the same samples (i.e., highly education female physicians, CEOs, etc.) in which observational studies have found what looks like attribution bias by gender.

### 3.6.4 Agents' Age

When we showed gender information about the agent, we also showed principals the age of the agents as a way to mask our research question. We chose age in

[^46]particular given the relatively small variation in age among university students. We can therefore check whether payments and beliefs are driven by this piece of information. When we run the same regressions as in Table 3.4 and 3.5, our results hold and the age coefficient is always consistently insignificantly different from 0 .

### 3.7 Discussion and Conclusion

Recent literature has suggested that a particular form of discrimination - attribution bias by gender - might affect assessments of actors' outcomes in economic environments differentially by gender. We conduct a pilot laboratory experiment in India to test for this effect. However, our results do not show evidence consistent with attribution bias by gender. While in our experiment principals' beliefs and payments are influenced by realized outcomes, we find no evidence that they differ by the agent's gender.

With the caveat that we have a relatively small sample size, our findings suggest that attribution bias by gender does not arise in a controlled environment. However, it is important to note that our findings need not imply that attribution bias by gender does not play a role at all. It is possible that in real-world settings, where gender is more salient, this bias naturally emerges.

## A Chapter 1 Appendix

## A. 1 Appendix A

## A.1.1 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.1.2 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.1.3 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.1.4 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.1.5 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.

## A. 2 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"

## A. 3 Comparative statics for theoretical framework

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

1. Comparative statics on $D$.

- $\frac{d c^{*}}{d D}=F^{-1^{\prime}}\left[\frac{-(1-p) u}{p v(D)}\right]\left[\left(\frac{-(1-p) u}{p}\right)\left(-\frac{v^{\prime}(D)}{(v(D))^{2}}\right)\right] \geq 0$
- $\frac{d q}{d D}=\frac{d c^{*}}{d D}+\frac{u v(D)+u D v^{\prime}(D)-(v(D))^{2}}{\left.(u-v(D))^{2}\right)}$
- Thus if;

$$
\begin{aligned}
& F^{-1^{\prime}}\left[\frac{-(1-p) u}{p v(D)}\right]\left[\left(\frac{-(1-p) u}{p}\right)\left(-\frac{v^{\prime}(D)}{(v(D))^{2}}\right)\right]+\frac{u v(D)+u D v^{\prime}(D)-(v(D))^{2}}{\left.(u-v(D))^{2}\right)} \leq 0 \\
& \text { then } \frac{d q}{d D} \leq 0
\end{aligned}
$$

Overall, sexual harassment is given by: $p\left(1-F\left(c^{*}\right)\right)$.
Therefore, $\frac{d S}{d c^{*}}=-p F^{\prime}\left(c^{*}\right) \frac{d c^{*}}{d D}$ and since we established $\frac{d c^{*}}{d D} \geq 0$, thus, $\frac{d S}{d c^{*}} \leq 0$
Relationships are given by: $R=q\left[p F\left(c^{*}\right)+(1-p)\right]$.
$\frac{d R}{d D}=\frac{d q}{d D}\left[p F\left(c^{*}\right)+(1-p)\right]+q p F^{\prime}\left(c^{*}\right) \frac{d c^{*}}{d D}$.
Relationships fall if :
$F^{-1^{\prime}}\left[\frac{-(1-p) u}{p v(D)}\right]\left[\left(\frac{-(1-p) u}{p}\right)\left(-\frac{1}{\left(v^{\prime}(D)\right)^{2}}\right)\right] \leq-\left[\frac{u\left(v(D)+u D v^{\prime}(D)-(v(D))^{2}\right.}{(u-v(D))^{2}}\right]\left[\frac{p F\left(c^{*}\right)+1-p}{p F\left(c^{*}\right)+(1-p)+q p F^{\prime}\left(c^{*}\right)}\right]$
2. Comparative statics on $p$.

- $\frac{d c^{*}}{d p}=F^{-1^{\prime}}\left[\frac{-(1-p) u}{p v(D)}\right]\left(\frac{u}{v(D) p^{2}}\right) \leq 0$ because $v(D)<0$
- $\frac{d q}{d p}=\frac{d c^{*}}{d p}$ and hence $\frac{d q}{d p}<0$.
$\frac{d S}{d p}=1-F\left(c^{*}\right)-p f\left(c^{*}\right) \frac{d c^{*}}{d p}>0$
$\frac{d R}{d p}=\frac{d q}{d p}\left[p F\left(c^{*}\right)+(1-p)\right]+q\left[F\left(c^{*}\right)+p f\left(c^{*}\right) \frac{d c^{*}}{d p}-1\right]$.
This can be rewritten as: $\frac{d R}{d p}=\frac{d q}{d p}\left[p F\left(c^{*}\right)+(1-p)\right]-q \frac{d S}{d p}<0$. This proves the result.


## A. 4 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 1.6.2)). 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). ${ }^{1}$ Lab-in-thefield 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.

[^47]
## A. 5 Attrition and Lee bounds, Female information, Difference in difference

Table A.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 $\times$ Female |  | 0.006 |
|  |  | $(0.032)$ |
|  |  |  |

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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* *}{ }^{*} \mathrm{p}<0.01$.


Figure A.1: 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 A.2: 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 A.3: 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 A.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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Sexual | Same Class | Mild | Intermediate | Extreme |
| :--- | :---: | :---: | :---: | :---: |
| Harassment | Index | events | events | 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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$. Asterisks denote significance: ${ }^{*} \mathrm{p}<0.1$, ${ }^{* *} \mathrm{p}<0.05$ and $^{* * *} \mathrm{p}<0.01$.

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

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

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: ${ }^{*} \mathrm{p}<0.1$, ${ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Control variable | $\beta_{1}$ | Control <br> 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. $\beta_{1}$ represents coefficient on treatment status of the woman. Robust standard errors are reported. *SC/ST/OBC represent castes in India.

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

| Control variable | Treatment <br> Mean | Control <br> 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 A.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 |
|  | $(0.025)$ | $(0.026)$ <br> Female <br> Female intervention $\times$ Female |
| N |  | $\left(0.0246^{* * *}\right.$ |
|  |  | -0.043 |
| $(0.036)$ |  |  |

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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

Table A.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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Sexual <br> Harassment | SH from <br> same class Index <br> (only extreme events sample) | SH from <br> same class Index <br> (All reported sample) | SH from <br> same class Index <br> (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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01 .^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and $^{* * *} \mathrm{p}<0.01$.

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

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

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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$. Asterisks denote significance: ${ }^{*}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

Table A.11: Sexual harassment from men outside college

| Sexual | Same Class <br> Harassment <br> Index <br> $(1)$ | Mild <br> events <br> $(2)$ | Intermediate <br> events <br> $(3)$ | Extreme <br> events <br> $(4)$ |
| :--- | :---: | :---: | :---: | :---: |
| 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.. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$. Asterisks denote significance: ${ }^{*}<0.1,{ }^{* *} \mathrm{p}<0.05$ and $^{* * *} \mathrm{p}<0.01$.

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

| Decision to <br> 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 $\times$ Male | 0.060 | 0.057 | $-0.137^{* *}$ | $-0.131^{* *}$ |
| treatment | $(0.062)$ | $(0.061)$ | $(0.058)$ | $(0.056)$ |
|  |  |  |  |  |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

|  |  | Survey Measures |  | Lab-in-Field |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
|  | Opposite sex <br> relationship <br> index | - | Dating | Opposite |  |
| sex friends |  |  |  |  |  | | Switch away |
| :---: |
| (same sex) | | Stick to |
| :---: |
| (opp sex) |

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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Dependent <br> variable | Perception of <br> legal costs index <br> $(1)$ | Perceived probability <br> of formal reporting <br> $(2)$ | Trust <br> ICC <br> $(3)$ |
| :--- | :---: | :---: | :---: |
| 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 |  |  |  |
| Control Mean | 1379 | 1377 | 1067 |

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. $P D S L A S S O$ is used for selecting controls. Asterisks denote significance: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Dependent <br> variable | Perception of <br> social costs <br> index <br> index <br> $(1)$ | Perceived <br> probability of <br> informal <br> reporting <br> $(2)$ | Perception of <br> peer support <br> against sexual <br> harassment <br> $(3)$ | Proportion men <br> nominated for <br> class mascot <br> class mascot <br> $(4)$ |
| :--- | :---: | :---: | :---: | :---: |
| 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 |
| Control Mean | 0.00 | 0.38 | 0.51 | 1129 |

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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| What would a woman from your class do <br> in hypothetical sexual harassment situations? | Situation 1 <br> $(1)$ | Situation 2 <br> $(2)$ | Situation 3 <br> $(3)$ |
| :--- | :---: | :---: | :---: |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

| Variable | Approached by man in same class <br> $(1)$ |
| :--- | :---: |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

Table A.19: Women's relationships with men outside the class
\(\left.$$
\begin{array}{lcc}\hline \hline & \text { Dating } & \begin{array}{c}\text { Opposite } \\
\text { (outside class) } \\
\text { (1) }\end{array}
$$ <br>
\hline \hline (outside class) <br>

(2)\end{array}\right]\)|  |  |  |
| :--- | :---: | :---: |
| All women |  |  |
| Male treatment | $\mathbf{0 . 0 6}{ }^{* * *}$ | -0.030 |
|  | $\mathbf{( 0 . 0 2 3 )}$ | $(0.062)$ |
| N |  |  |
| Control mean | 0.146 | 1,347 |

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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

|  | Situation 1 <br> $(1)$ | Situation 2 <br> $(2)$ | Situation 3 <br> $(3)$ |
| :--- | :---: | :---: | :---: |
| Men's beliefs about male classmates |  |  |  |
| Male treatment | $0.0659^{* * *}$ <br> $(0.0245)$ | $0.0611^{* * *}$ <br> $(0.0142)$ | $0.0422^{* *}$ <br>  <br> N |
| Control Mean | 1310 | 1310 | 1310 |
| Men's beliefs about female classmates |  |  | 0.38 |
| Male treatment | 0.52 | 0.35 |  |
|  | $\left(0.0739^{* * *}\right.$ | $0.0950)$ | $(0.0194)$ |
| N | 1310 | 1310 | $(0.0201)$ |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

|  | Situation 1 <br> $(1)$ | Situation 2 <br> $(2)$ | Situation 3 <br> $(3)$ |
| :--- | :---: | :---: | :---: |
| 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: ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$.

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

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

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. ${ }^{*} \mathrm{p}<0.1,{ }^{* *} \mathrm{p}<0.05$ and ${ }^{* * *} \mathrm{p}<0.01$. Asterisks denote significance: ${ }^{*}<0.1,{ }^{* *} \mathrm{p}<0.05$ and $^{* * *} \mathrm{p}<0.01$.

## B Chapter 2 Appendix

## B. 1 Appendix B

## B.1.1 Screenshots of the Experiment

Here we provide the screenshots for T2. Screenshots for the other experimental treatments are available upon request.

## B.1.2 Welcome Page

Figure B.1: Screenshot of the Welcome Page Template

## Welcome to this study!

Welcome to this study and thanks for your participation.
In this experiment you will earn a show-up fee of $£ 5.00$. You can earn extra money depending on the decisions that you make during the experiment. In particular, the experiment will consist of some decision tasks. In each of these tasks you can earn extra money (unless stated otherwise). However, at the end of the experiment only one of these tasks will be selected to actually count for your payments. Each decision task is equally likely to count for payments.

After you have completed these tasks, you will be asked to answer some survey questions. For this part you will earn some money for sure and some extra money that will depend on your decisions.

All your decisions in the experiment will be anonymous. In particular, neither participants nor the experimenter can match your decisions to your identity.

Please notice that our experiment does not entail any sort of deception, as all experiments conducted in economics. This implies that you will be given truthful information regarding the instructions of the experiment and the experimental tasks.

Please note that no communication is allowed during the experiment. You are also not allowed to use mobile phones or surf the Internet. Violation of these rules will result in the exclusion from the experiment and all payments.

If you have any questions during the experiment, please raise your hand and the experimenter will come to you and answer them in private.

## B.1.3 Instructions - the Quiz

Figure B.2: Screenshot of the Instructions Template for the Quiz

## Instructions - The Quiz (Decision Task I)

In this task you are asked to solve a quiz. Specifically, you will be asked 25 questions, and each question will test your aptitude in one of the following five categories: Assembling Objects, Sports, Mechanical Comprehension, Math Knowledge, and General Science.

You will have 10 minutes to answer as many questions as possible. All questions will appear on the same page and the computer will keep time for you. You may answer the questions in any order.

If this task is selected to count for payments, you will be paid for three randomly chosen questions. For each correct question (out of these three) you will earn $£ 2.00$. Thus, in total you can earn up to $£ 6.00$ in this task.

On the next page, you will be asked some comprehension questions about the instructions. You can only proceed with the experiment if you have solved them correctly.

To continue with the comprehension questions for this task, please type in the cell below the number " 100 ".
$\square$

## B.1.4 Comprehension Questionnaire - the Quiz

Figure B.3: Screenshot of the Comprehension Questionnaire Template for the Quiz

## Decision Task 1 - Comprehension Questions

Please answer below the comprehension questions about this task. You will not be able to proceed until you answer all of them correctly.

If you find any difficulties answering the questions please refer back to the instructions located below the "next" button.

If you have any questions or doubts, please raise your hand.

How many questions do you have to solve?
--------- V

How much time do you have in total to solve the questions?

$\checkmark$

How much money would you earn if you answer THREE questions correctly out of the three randomly chosen questions?
$\qquad$

Do you lose money for not answering correctly one of these three randomly chosen questions?
$\qquad$ $\checkmark$

[^48]
## B.1.5 The Quiz

Figure B.4: Screenshot of the Quiz

## Decision Task 1 - The Quiz

Time left to complete this page: 8:57

Please, answer the questions below. You have 10 minutes ( 600 seconds) to answer as many questions as possible.

## Category: Assembling Objects

Question 1. Which figure best shows how the objects in the left box will touch if the letters for each object are matched?


## Question 1 Answer:



Question 2. Which figure best shows how the objects in the left box will appear if they are fit together?

## B.1.6 Instructions - Belief Elicitation

Figure B.5: Screenshot of the Instructions Template for the Prior Belief Elicitation

## Instructions - Decision Task 2

In the next screen, you will be asked to assess your own performance in the Quiz as compared to other people's performance in this lab session.

That is, your task is to guess whether your performance (the number of correctly solved questions) in the Quiz is in the top half of the distribution of these participants. Ties are broken randomly. In particular, we will ask you to state the probability with which you think that your score is in the top half (that is top $50 \%$ ) among these participants.

Specifically, your score will be compared to all other participants' performances in today's experiment, except for the score of one randomly chosen participant. The score of this participant will not be part of the distribution.

If this task is selected for payment, you have the opportunity to earn $£ 2.00$ or $£ 6.00$ depending on your reported belief about you being in the top half and whether you are actually in the top half.

The payment mechanism is such that reporting your true belief about your performance will maximize your chance of earning the higher payment of $£ 6.00$. (If you would like to understand why it is in your best interest to tell us your true belief, please click on the button "Detailed Explanation" that you can find below the next button).

To continue please type in the cell below the number " 30 ".

## B.1.7 Prior Belief Elicitation

Figure B.6: Screenshot of the Prior Belief Elicitation

## Your Rank in the Distribution - Your Decision

By adjusting the slider below, please state the probability with which you think that you scored in the top half of the distribution (that is, as compared to other people who have completed the same task as you).

The initial position of the slider is randomly determined (it is NOT related to your actual rank).

Please adjust the slider to state the probability with which you think that you are in the top half of the distribution.

## B.1.8 Instructions - Information Structure Choice (T1)

Figure B.7: Screenshot of the Instructions Template for the Information Structure Choice - Part I

## Instructions Task 2 - The Game

For this part of the experiment you will be paired with one random participant in this lab session. In particular, each one of you two will be making two similar decisions which may or may not affect each other.

## Feedback

Depending on your rank in the distribution of performances, you will receive one feedback (that is, one piece of information) about your rank. You can receive three types of feedback in the form of evaluations:

- The green ball that tells you: "TOP HALF";
- The red ball that tells you: "BOTTOM HALF";
- Furthermore, you can also receive a blue "NO EVALUATION" feedback.

The figure below shows you the exact three possible evaluations that you can receive.


## How is the feedback determined?

Which feedback you receive depends on your actual rank in the distribution in the quiz and the evaluation system from which the feedback is generated.

In particular, there are two possible evaluation systems and you will have to choose from which evaluation system you want to receive feedback about your performance.

## Why it matters?

Following the feedback, we will ask you to state again your belief about the probability with which you think that you are in the top half of the distribution and your answer to this question will be paid for. Thus, the feedback you receive can help you in answering that question and, thus, in increasing your payments in case that question is randomly selected for payments.

Figure B.8: Screenshot of the Instructions Template for the Information Structure Choice - Part II

The different components of the evaluation systems are as follows:

1. Precision of evaluation

You will always receive your feedback with some probability of error. This probability of error depends on the precision of the evaluation system. In particular, the precision of the Public evaluation system is $90 \%$, while that of the Private evaluation system is $60 \%$. This means that the information regarding your rank that you receive from the Public evaluation system is correct $90 \%$ of the times, while it is correct $60 \%$ of times for the private evaluation system.

To better understand how the precision works, please refer to the picture below. If you are in the top half and you choose the public evaluation system, then you will receive the green (correct) evaluation 9 out of 10 times. Similarly, if you are in the bottom half, you will receive the red (correct) evaluation 9 out of 10 times. Thus, in each case you will get the wrong evaluation 1 out of 10 times. If you instead choose the private evaluation system and you are in the top half then you will receive the green (correct) evaluation 6 out of 10 times. Similarly, if you are in the bottom half, you will receive the red (correct) evaluation 6 out of 10 times. Thus, in each case you will get the wrong evaluation 4 out of 10 times.

Public Evaluation System
 Bottom Half

Private Evaluation System


Figure B.9: Screenshot of the Instructions Template for the Information Structure Choice - Part III

## 2. Mode of evaluation

In the Public evaluation system your evaluation will also be shown to the participant you are paired with in this task. That is, if you get a green evaluation, your partner will also get to see this evaluation about your performance. Similarly, if you get a red evaluation, your partner will get to see it too. In the private evaluation system, instead, the evaluation will be privately communicated to you. Therefore, the participant you are paired with will not get to see your evaluation. In fact, regardless of the evaluation that you receive, he/she will see the following picture:

## NO EVALUATION

## Your partner's feedback

mportantly, while you are making all these choices, the participant you are matched with will be making the same choices. Thus, if he/she chooses the Public evaluation system, then you will get to see his/her evaluation. If, instead, he/she opts for the Private evaluation system, then you will not see his/her feedback.

## Summary

In sum, the Public evaluation system is more informative about your actual performance than the Private evaluation system since it has higher precision. Also, in contrast to the Private evaluation system, your paired participant gets to see your feedback in the public evaluation system (i.e., the evaluation you receive).

Similarly, the chosen evaluation system by your partner will determine whether you see his/her feedback.

To continue with the comprehension questions for this part of task 2, please type in the cell below the number "40".
$\square$

## B.1.9 Comprehension Questions - Information Structure Choice

Figure B.10: Screenshot of the Comprehension Questionnaire Template for the Information Structure Choice

## The Game - Comprehension Questions

Please answer below the comprehension questions about this task.
You will not be able to proceed until you answer all of them correctly.
If you find any difficulties answering the questions please refer back to the instructions located below the "next" button.

If you have any questions or doubts, please raise your hand.

1. What determines your feedback?

2. What is the probability that you get a green evaluation if you are in the top half, for each evaluation system?

3. What is the probability that you get a red evaluation if you are in the bottom half, for each evaluation system?

4. If you choose the PRIVATE evaluation system, what will your partner see about your performance?
$\qquad$
5. If you choose the PUBLIC evaluation system, what will your partner see about your performance?


## B.1.10 Information Structure Choice

Figure B.11: Screenshot of the Information Structure Choice

## Decision about the Evaluation System

Now, please choose whether you want to receive feedback from the Public or Private evaluation systems.
Please find below the visual description of the two evaluation systems again.
Please remember that your partner also gets to see your feedback in case you choose the Public evaluation system. While this is not the case if you choose the Private evaluation system.

Public Evaluation System


Private Evaluation System


Your Choice:
$\qquad$

## B.1.11 Partner's Feedback

Figure B.12: Screenshot of the Feedback the Subject Received about her Partner

## Partner's Feedback and your Choice

## Feedback of Your Partner

Your partner has chosen to receive feedback about his/her performance from the Private Evaluation System.
Your partner's feedback about his/her performance in the Quiz is the following:

## NO EVALUATION

By adjusting the slider below, please state the probability with which you think that your partner scored in the top half of the distribution. This question is not monetarily incentivised.

The initial position of the slider is randomly determined (it is NOT related to your partner's actual rank).

Please adjust the slider to state the probability with which you think that your partner is in the top half of the distribution.


Next


## B.1.12 Posterior Belief

Figure B.13: Screenshot of the Feedback Received and Elicitation of Posterior Belief

## Your Rank in the Distribution - Your Decision

```
Your previous guess, that you are in the top half of the distribution in the Quiz, was 41 percent.
```

Please, find below the feedback that you received from the Private evaluation system on your performance .

## TOP HALF

By adjusting the slider below, please state the probability with which you think that you scored in the top half of the distribution (that is, as compared to other people who have completed the same task as you).

Specifically, and just as before, your score was compared to all other participants' performances in today's experiment, except for the score of one randomly chosen participant (who is your partner). The score of this participant was not part of the distribution.

Remember that the payment mechanism is such that reporting your true belief about your performance will maximize your chance of earning the higher payment of $£ 6.00$. (Again, if you would like to remember why it is in your best interest to tell us your true belief, please click on the button "Detailed Explanation" that you can find below the next button).

Please adjust the slider to state the probability with which you think that you are in the top half of the distribution.

## B.1.13 Instructions for strategic treatment (T4)

Figure B.14: Screenshot of information about payments for Strategic Treatment

## Instructions Task 2 - The Partnering Decision

Before you see your own evaluation you will get to see the feedback of the person you are paired with. If they chose to go for public evaluation then you will be shown their correct feedback with $80 \%$ probability, wrong feedback for them with $10 \%$ probability and no signal or feedback with $10 \%$ probability. If they chose to go with private feedback then you will see no signal as well as below.

The figure below shows you the exact three possible evaluations that you will see for your paired participant


## B.1.14 Instructions for strategic treatment (T4) continued

Figure B.15: Screenshot of information about payments for Strategic Treatment

The figure below shows you what kind of feedback you will see depending on the evaluation system chosen by your paired participant.

Partner chooses public feedback Partner chooses private feedback


Once you get to see their feedback, you have to make the decision on whether you would like to partner with them. On the basis of your decision to partner with your paired participant will affect your payoffs in the following manner in this part of the experiment. In particular:

- If you decide to partner with them and they are really in the top half of the distribution, then you will be paid $£ 6$.
- If you decide to partner with them and they are really in the bottom half of the distribution, then you will be paid $\mathbf{0}$.
- If you decide to not partner with them and instead go solo, then you will be paid $£ 2.5$.


## B.1.15 Instructions for strategic treatment (T4) continued

Figure B.16: Screenshot of information about payments for Strategic Treatment
Thus, it is in your best economic interest to partner with him/her if you think your paired participant is in the top half, whereas it is best for you not to partner with him/her if he/she is in the bottom half on the basis of the feedback you see for them. Please note whether you get to see their feedback or not depends on whether you paired participant chose for public or private feedback.

Importantly, the score of your randomly matched participant was not part of the distribution of scores your performance has been compared to. This implies that whether he/she is in the top half does not imply that you are less likely to be in the top half. Similarly, if he/she is in the bottom half, you are not more likely to be in the top half.

## Your paired participant's Hiring Decision

Similarly, the participant you are paired with will also make a partnering decision about you. And before he/she makes the decision, they will also get to see your evaluation if you chose the public evaluation system. They will also see no feedback if you chose private evaluation system. This is to help them take the decision for partnering with you in the same way as we show you their feedback to help you make the same partnering decision.

Importantly, if your paired participant decides to partner you, you will earn $£ 2.5$ irrespective of your rank (and if this part of the task is randomly selected to count for you earnings). If not, you will earn 0 . Please note though, your chances of being chosen for partnering with them are better if they think that you are in the top half of distribution. Their partnering incentives are just like yours.

Thus, it is in your best economic interest that your paired participant decides to "partner" with you.
The following are your payoffs if your paired participant decides to partner with you

| Your decision to partner or not with them | Your partner is truly in the top half | Your partner is truly in the bottom half |
| ---: | :---: | :---: |
| Partner with the paired participant | $£ 6$ | 0 |
| Do not partner with the paired participant | $£ 2.5$ | $£ 2.5$ |

## B.1.16 Distribution of questions answered correctly

Figure B.17: Distribution of number of questions answered correctly by treatment.


Notes: The graph provides the distribution of the number of questions that subjects answered correctly by the treatment they were assigned to.

## B.1.17 Qualitative answers

Figure B.18: Qualitative answers to reasons behind choosing less informative feedback.


Notes: The graph provides the different reasons for which men and women chose less informative feedback out of those who chose less informative feedback.

## B.1.18 Efficiency versus equity

Figure B.19: Efficiency versus equity.


Notes: The graph provides percentage of men and women who take-up less informative feedback from.

## B. 2 Summary Statistics

Table B.1: Descriptive statistics

|  | T 0 | T 1 | T 2 | T 3 |
| :--- | :---: | :---: | :---: | :---: |
|  |  |  |  |  |
| Female (Share) | 0.576 | 0.616 | 0.520 | 0.544 |
|  | $(0.498)$ | $(0.489)$ | $(0.502)$ | $(0.501)$ |
| Age (Mean) |  |  |  |  |
|  | 21.591 | 21.279 | 21.147 | 21.656 |
| Quantitative Education (Share) | $(3.658)$ | $(4.881)$ | $(4.225)$ | $(4.117)$ |
|  | 0.697 | 0.616 | 0.637 | 0.633 |
|  | $(0.463)$ | $(0.489)$ | $(0.483)$ | $(0.485)$ |
| Risk Preferences (Mean) | 5.939 | 6.151 | 5.657 | 5.767 |
|  | $(2.089)$ | $(1.863)$ | $(1.832)$ | $(1.949)$ |
| Risk choice (Mean) | 48.136 | 50.988 | 51.794 | 45.011 |
|  | $(31.823)$ | $(30.913)$ | $(29.707)$ | $(30.593)$ |
| Score (Mean) | 12.879 | 13.279 | 13.549 | 13.378 |
|  | $(2.726)$ | $(2.965)$ | $(3.053)$ | $(3.136)$ |
| Prior (Mean) | 54.773 | 51.267 | 53.373 | 53.533 |
|  | $(21.651)$ | $(21.982)$ | $(22.517)$ | $(22.821)$ |
| Experience with experiments (Share) | 0.500 | 0.523 | 0.500 | 0.467 |
|  | $(0.504)$ | $(0.502)$ | $(0.502)$ | $(0.502)$ |
| Friends in the experiment (Share) | 0.606 | 0.221 | 0.275 | 0.167 |
|  | $(0.492)$ | $(0.417)$ | $(0.448)$ | $(0.375)$ |
| N |  | 66 | 86 | 102 |

Notes: the table shows descriptive statistics (in means or in shares) of our experimental subjects by treatment. Female is the share of female subjects. Age is the reported age of the subject. Quantitative education is a dummy variable equal to 1 if the subject's course of study is mainly quantitative. Risk preferences and Risk choice are two variables that capture subjects' risk preferences. The first is the Dohmen et al. (2011) risk elicitation question, while the second is the Gneezy and Potters (1997) risky choice question. Score is the subject's score in the quiz and Prior is the subject's prior belief about her relative performance in the quiz. Experience with experiments is a dummy variable that equals 1 if the subject has taken part to at least 3 other experiments in the past. The Friends in the experiment dummy variable is equal to 1 if the subject knows at least 1 other participant in her session. Standard deviations are in parentheses.

## B.2.1 Demographic questions

The questions we asked are:

- What is your gender?( Male, Female, Other)
- What is your age? (Text answer)
- What is your mother tongue? (Text answer)
- What is your country of origin? (Text answer)
- What is your ethnicity? (Text answer)
- What is your year of study? (1st, 2nd, 3rd, 4th, +5 years, Other (I am not a student))
- In which department are you currently enrolled? Text answer
- Have you taken part in experiments before? $(0,1,2,3,4,+4$ experiments $)$
- Do you know one or more participants in the experiment today? $(0,1,2,3,4,+4$ participants)
- How willing are you to take risks? (0 (not willing at all), $1,2,3,4,5,6,7,8,9,10$ (very willing))

Table B.2: Descriptive statistics by information structure chosen

|  | Information structure A | Information structure B |
| :--- | :---: | :---: |
| Female (Share) | 0.534 | 0.696 |
|  | $(0.499)$ | $(0.463)$ |
| Age (Mean) | 21.336 | 21.714 |
|  | $(4.35)$ | $(3.77)$ |
| Year of study | 1.576 | 1.517 |
|  | $(0.494)$ | $(0.504)$ |
| Quantitative Education (Share) | 0.697 | 0.616 |
|  | $(0.463)$ | $(0.489)$ |
| Risk Preferences (Mean) | 5.861 | 5.875 |
|  | $(1.909)$ | $(2.010)$ |
| Risk choice (Mean) | 49.174 | 48.820 |
|  | $(31.188)$ | $(27.930)$ |
| Score (Mean) | 12.879 | 13.279 |
|  | $(2.726)$ | $(2.965)$ |
| Prior (Mean) | 53.701 | 50.357 |
|  | $(22.900)$ | $(23.274)$ |
| Experience with experiments (Share) | 1.50 | 1.48 |
|  | $(0.501)$ | $(0.504)$ |
| Friends in the experiment (Share) | 0.295 | 0.303 |
|  | $(0.456)$ | $(0.463)$ |
| N | 288 | 56 |

Notes: the table shows descriptive statistics (in means or in shares) of our experimental subjects by the information structure they chose to get the feedback from for all treatments together. Female is the share of female subjects. Age is the reported age of the subject. Quantitative education is a dummy variable equal to 1 if the subject's course of study is mainly quantitative, year of study is the year of undergraduate study they are in if they are an undergraduate. Risk preferences and Risk choice are two variables that capture subjects' risk preferences. The first is the Dohmen et al. (2011) risk elicitation question, while the second is the Gneezy and Potters (1997) risky choice question. Score is the subject's score in the quiz and Prior is the subject's prior belief about her relative performance in the quiz. Experience with experiments is a dummy variable that equals 1 if the subject has taken part to at least 3 other experiments in the past. The Friends in the experiment dummy variable is equal to 1 if the subject knows at least 1 other participant in her session. Standard deviations are in parentheses.

Table B.3: Choice of Feedback Mode B in (T1) with session fixed effects

|  | OLS |  |  |
| :--- | :--- | :--- | :--- |
|  | $(1)$ <br> Info B | $(2)$ <br> Info B | $(3)$ <br> Info B |
|  |  |  |  |
| Female | $0.164^{*}$ | $0.182^{*}$ | $0.217^{*}$ |
|  | $(0.092)$ | $(0.098)$ | $(0.127)$ |
|  |  |  |  |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |
| Session FE | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.098 | 0.117 | 0.214 |
| N | 66 | 66 | 66 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue, country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.

Table B.4: Choice of Information Structure B in (T1)-(T3) with session FE

|  | OLS |  |  |
| :--- | :--- | :--- | :--- |
|  | $(1)$ <br> Info B | $(2)$ <br> Info B | $(3)$ <br> Info B |
| Female | 0.1644 <br> $(0.0918)^{*}$ | 0.1599 <br> $(0.0935)^{*}$ | 0.1724 |
| $(0.0951)^{*}$ |  |  |  |
| T2 | 0.0195 | 0.0160 | 0.1118 |
|  | $(0.1139)$ | $(0.1148)$ | $(0.1231)$ |
|  | 0.1150 | 0.1131 | 0.1201 |
| T3 | $(0.1425)$ | $(0.1433)$ | $(0.1469)$ |
|  |  |  |  |
| T2 $\times$ Female | -0.0659 | -0.0713 | -0.1205 |
|  | $(0.1117)$ | $(0.1104)$ | $(0.1120)$ |
| T3 $\times$ Female | -0.0924 | -0.0967 | -0.1162 |
|  | $(0.1249)$ | $(0.1242)$ | $(0.1212)$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |
| Session FE | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.0671 | 0.0682 | 0.1323 |
| N | 254 | 254 | 254 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.

Table B.5: Overall difference in take up of less informative feedback, all treatments with session fixed effects

|  | OLS |  |  |
| :--- | :--- | :--- | :--- |
|  | $(1)$ <br> Info B | $(2)$ <br> Info B | $(3)$ <br> Info B |
|  |  |  |  |
| Female | $0.1644^{*}$ | $0.1608^{*}$ | $0.1715^{*}$ |
|  | $(0.0917)$ | $(0.0930)$ | $(0.0949)$ |
| T2 | 0.0195 | 0.0167 | 0.0720 |
|  | $(0.1139)$ | $(0.1145)$ | $(0.1185)$ |
| T3 |  |  |  |
|  | 0.1150 | 0.1135 | 0.1127 |
|  | $(0.1424)$ | $(0.1430)$ | $(0.1523)$ |
| T4 | 0.1229 | 0.1257 | 0.1476 |
|  | $(0.1380)$ | $(0.1384)$ | $(0.1457)$ |
| T2 $\times$ Female | -0.0659 | -0.0702 | -0.1036 |
|  | $(0.1116)$ | $(0.1106)$ | $(0.1123)$ |
| T3 $\times$ Female | -0.0924 | -0.0958 | -0.1216 |
|  | $(0.1248)$ | $(0.1243)$ | $(0.1236)$ |
| T4 $\times$ Female | -0.1830 | -0.1886 | $-0.2299^{*}$ |
|  | $(0.1291)$ | $(0.1290)$ | $(0.1318)$ |
| Prior Belief |  |  |  |
| Demographics <br> Session FE | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| P-value $\beta_{1}+\beta_{7}=0$ | 0.837 | 0.765 | 0.537 |
| R-Squared | 0.057 | 0.058 | 0.099 |
| N | 344 | 344 | 344 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Robust standard errors are reported in parentheses.

Table B.6: Choice of Feedback Mode B in (T1) with standard errors clustered at the session level

|  | OLS |  |  | Probit |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | (1) <br> Info B | (2) Info B | (3) <br> Info B | (4) <br> Info B | (5) <br> Info B | (6) <br> Info B |
| Female | $\begin{aligned} & 0.2011^{*} \\ & (0.0851) \end{aligned}$ | $\begin{aligned} & 0.2202^{*} \\ & (0.0891) \end{aligned}$ | $\begin{aligned} & 0.2387 \\ & (0.1369) \end{aligned}$ | $\begin{aligned} & 0.201^{* *} \\ & (0.084) \end{aligned}$ | $\begin{aligned} & 0.210^{* * *} \\ & (0.081) \end{aligned}$ | $\begin{aligned} & 0.175^{*} \\ & (0.095) \end{aligned}$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |
| WC bootstrap p values | 0.0625 | 0.0625 | 0.1875 | 0.0625 | 0.0625 | 0.0625 |
| R-Squared | 0.077 | 0.098 | 0.205 | 0.105 | 0.131 | 0.313 |
| N | 66 | 66 | 66 | 66 | 66 | 66 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue,country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors are clustered at the session level. p-values from wild cluster bootstrap is reported in the table.

Table B.7: Choice of Information Structure B in (T1)-(T3) with clustered standard errors

|  | OLS |  |  | Probit |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} (1) \\ \text { Info B } \end{gathered}$ | $\begin{gathered} (2) \\ \text { Info B } \end{gathered}$ | (3) <br> Info B | $\begin{gathered} (4) \\ \text { Info B } \end{gathered}$ | (5) <br> Info B | (6) <br> Info B |
| Female | $\begin{aligned} & 0.201^{* *} \\ & (0.079) \end{aligned}$ | $\begin{aligned} & 0.197^{* *} \\ & (0.080) \end{aligned}$ | $\begin{aligned} & 0.197^{* *} \\ & (0.088) \end{aligned}$ | $\begin{aligned} & 0.213^{* *} \\ & (0.010) \end{aligned}$ | $\begin{aligned} & 0.210^{* *} \\ & (0.104) \end{aligned}$ | $\begin{aligned} & 0.203^{* *} \\ & (0.093) \end{aligned}$ |
| T2 | $\begin{aligned} & 0.025 \\ & (0.050) \end{aligned}$ | $\begin{aligned} & 0.026 \\ & (0.050) \end{aligned}$ | $\begin{aligned} & 0.066 \\ & (0.047) \end{aligned}$ | $\begin{aligned} & 0.055 \\ & (0.123) \end{aligned}$ | $\begin{aligned} & 0.057 \\ & (0.123) \end{aligned}$ | $\begin{aligned} & 0.114 \\ & (0.108) \end{aligned}$ |
| T3 | $\begin{aligned} & 0.066 \\ & (0.058) \end{aligned}$ | $\begin{aligned} & 0.067 \\ & (0.059) \end{aligned}$ | $\begin{aligned} & 0.092 \\ & (0.070) \end{aligned}$ | $\begin{aligned} & 0.118 \\ & (0.125) \end{aligned}$ | $\begin{aligned} & 0.121 \\ & (0.125) \end{aligned}$ | $\begin{aligned} & 0.149 \\ & (0.118) \end{aligned}$ |
| $\mathrm{T} 2 \times$ Female | $\begin{aligned} & -0.092 \\ & (0.089) \end{aligned}$ | $\begin{aligned} & -0.097 \\ & (0.088) \end{aligned}$ | $\begin{aligned} & -0.126 \\ & (0.093) \end{aligned}$ | $\begin{aligned} & -0.087 \\ & (0.093) \end{aligned}$ | $\begin{aligned} & -0.092 \\ & (0.091) \end{aligned}$ | $\begin{aligned} & -0.097 \\ & (0.060) \end{aligned}$ |
| $\mathrm{T} 3 \times$ Female | $\begin{aligned} & -0.114 \\ & (0.099) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.112 \\ & (0.101) \end{aligned}$ | $\begin{aligned} & -0.115 \\ & (0.099) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.081) \end{aligned}$ | $\begin{aligned} & -0.096 \\ & (0.061) \end{aligned}$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |
| WC bootstrap p-values | 0.10 | 0.12 | 0.13 | 0.0781 | 0.100 | 0.085 |
| R-Squared | 0.036 | 0.037 | 0.104 | 0.049 | 0.049 | 0.160 |
| N | 254 | 254 | 254 | 254 | 254 | 254 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include:age, mother tongue, country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors are clustered at the session level and reported in parentheses. Wild cluster bootstrap p values are reported as well.

Table B.8: Overall difference in take-up of less informative feedback, all treatments with clustered standard errors.

|  | OLS |  |  | Probit |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} (1) \\ \text { Info B } \end{gathered}$ | (2) Info B | $(3)$ Info B | $\begin{gathered} (4) \\ \text { Info B } \end{gathered}$ | $\begin{gathered} (5) \\ \text { Info B } \end{gathered}$ | (6) <br> Info B |
| Female | $\begin{aligned} & 0.201^{* *} \\ & (0.078) \end{aligned}$ | $\begin{aligned} & 0.198^{* *} \\ & (0.080) \end{aligned}$ | $\begin{aligned} & 0.2011^{* *} \\ & (0.089) \end{aligned}$ | $\begin{aligned} & 0.237^{* *} \\ & (0.115) \end{aligned}$ | $\begin{aligned} & 0.235^{* *} \\ & (0.115) \end{aligned}$ | $\begin{aligned} & 0.242^{* *} \\ & (0.111) \end{aligned}$ |
| T2 | $\begin{aligned} & 0.025 \\ & (0.049) \end{aligned}$ | $\begin{aligned} & 0.026 \\ & (0.049) \end{aligned}$ | $\begin{aligned} & 0.040 \\ & (0.050) \end{aligned}$ | $\begin{aligned} & 0.062 \\ & (0.141) \end{aligned}$ | $\begin{aligned} & 0.064 \\ & (0.142) \end{aligned}$ | $\begin{aligned} & 0.088 \\ & (0.132) \end{aligned}$ |
| T3 | $\begin{aligned} & 0.066 \\ & (0.058) \end{aligned}$ | $\begin{aligned} & 0.067 \\ & (0.058) \end{aligned}$ | $\begin{aligned} & 0.082 \\ & (0.066) \end{aligned}$ | $\begin{aligned} & 0.138 \\ & (0.149) \end{aligned}$ | $\begin{aligned} & 0.139 \\ & (0.150) \end{aligned}$ | $\begin{aligned} & 0.175 \\ & (0.149) \end{aligned}$ |
| T4 | $\begin{aligned} & 0.184^{* * *} \\ & (0.044) \end{aligned}$ | $\begin{aligned} & 0.186^{* * *} \\ & (0.045) \end{aligned}$ | $\begin{aligned} & 0.198^{* * *} \\ & (0.052) \end{aligned}$ | $\begin{aligned} & 0.294^{*} \\ & (0.159) \end{aligned}$ | $\begin{aligned} & 0.299^{* *} \\ & (0.152) \end{aligned}$ | $\begin{aligned} & 0.356^{* *} \\ & (0.148) \end{aligned}$ |
| T2 $\times$ Female | $\begin{aligned} & -0.092 \\ & (0.103) \end{aligned}$ | $\begin{aligned} & -0.096 \\ & (0.102) \end{aligned}$ | $\begin{aligned} & -0.118 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.094 \\ & (0.099) \end{aligned}$ | $\begin{aligned} & -0.098 \\ & (0.096) \end{aligned}$ | $\begin{aligned} & -0.111 \\ & (0.073) \end{aligned}$ |
| T3 $\times$ Female | $\begin{aligned} & -0.114 \\ & (0.105) \end{aligned}$ | $\begin{aligned} & -0.117 \\ & (0.104) \end{aligned}$ | $\begin{aligned} & -0.110 \\ & (0.104) \end{aligned}$ | $\begin{gathered} -0.124^{*} \\ (0.082) \end{gathered}$ | $\begin{aligned} & -0.127 \\ & (0.081) \end{aligned}$ | $\begin{aligned} & -0.126^{* *} \\ & (0.065) \end{aligned}$ |
| T4 $\times$ Female | $\begin{gathered} -0.196^{*} \\ (0.119) \end{gathered}$ | $\begin{aligned} & -0.202^{*} \\ & (0.118) \end{aligned}$ | $\begin{gathered} -0.227^{*} \\ (0.121) \end{gathered}$ | $\begin{aligned} & -0.162^{* * *} \\ & (0.051) \end{aligned}$ | $\begin{aligned} & -0.164^{* * *} \\ & (0.049) \end{aligned}$ | $\begin{aligned} & -0.173^{* * *} \\ & (0.038) \end{aligned}$ |
| Prior Belief |  | $\checkmark$ | $\checkmark$ |  | $\checkmark$ | $\checkmark$ |
| Demographics |  |  | $\checkmark$ |  |  | $\checkmark$ |
| WC bootstrap p-values |  |  |  |  |  |  |
| $\beta_{1}$ | 0.07 | 0.09 | 0.16 | 0.07 | 0.10 | 0.11 |
| $\beta_{4}$ | 0.10 | 0.11 | 0.11 | 0.07 | 0.06 | 0.05 |
| $\beta_{7}$ | 0.03 | 0.026 | 0.02 | 0.02 | 0.02 | 0.01 |
| $\beta_{1}+\beta_{7}$ | 0.90 | 0.96 | 0.71 | 0.92 | 0.94 | 0.67 |
| P -value $\beta_{1}+\beta_{7}=0$ | 0.932 | 0.948 | 0.726 | 0.936 | 0.916 | 0.626 |
| R-Squared | 0.036 | 0.036 | 0.365 | 0.042 | 0.043 | 0.103 |
| N | 344 | 344 | 344 | 344 | 344 | 344 |

The dependent variable is a dummy equal to 1 if the participant chose the less informative information structure (Info B) and 0 otherwise (Info A). Demographic variables include: age, mother tongue, country of origin, ethnicity, education, department of study, risk preferences, prior experience with experiments and whether they know anyone else in their experimental session or not. Standard errors clustered at the session level are reported in parentheses. Wild cluster bootstrap p values are added for the key coefficient estimates.

## C Chapter 3 Appendix

## C. 1 Appendix C: Summary Statistics

## C.1.1 Mean Payments to Each Party by Realized Outcome

In the paper we have analysed principals' payment decisions to their matched agent following low and high outcomes. Here, we now present a visual representation of average payments to each party following both low and high outcomes (see Figure C.1). This figure shows that, going from a low to a high outcome, principals' payments to their matched agent increase (from ₹ 135.15 to ₹ 243.35 ) whereas it decrease for both payments to the other randomly drawn agent (from ₹ 120.37 to ₹ 75.99 ) and to the experimenter (from ₹ 94.48 to ₹ 30.66 ).

## C.1.2 Mean Payments to Each Party by Realized Outcome and Gender of the Agent

If we look at mean payments by taking into account the gender of the matched agent, we find very similar patterns (Figure C.2). Indeed, while agents (irrespective of their gender) are being rewarded for high outcomes, this comes at the cost of lower payments to both the other randomly matched agent and the experimenter.

## C.1.3 Robustness Checks for Principals' Payment Decisions and Beliefs

Table C.1: Summary statistics of our sample

|  | Principals | Agents |
| :--- | :---: | :---: |
| Female | $73 \%$ | $55 \%$ |
|  | $(0.44)$ | $(0.50)$ |
| Age | 22.92 | 22.02 |
|  | $(1.18)$ | $(1.18)$ |
| Degree of study | 3.21 | 3.21 |
|  | $(1.02)$ | $(1.02)$ |
| Year of study | 2.21 | 2.14 |
|  | $(0.95)$ | $(0.84)$ |
| Language | 1.98 | 2.00 |
|  | $(0.15)$ | $(0.00)$ |
| Religion | 2.12 | 2.07 |
|  | $(1.02)$ | $(0.51)$ |
| Caste | 3.52 | 3.36 |
|  | $(0.83)$ | $(0.98)$ |
| N | 42 | 42 |

Notes: Table shows descriptive statistics (in means) of the experimental dataset. Standard deviations are in parentheses. Female is the share of female participants. Age is the reported age of the participant. Degree of study: $1=$ Sociology, $2=$ Commerce, $3=$ Geography, $4=$ Economics, $5=$ Other. Year of study: $1=$ First year master degree, $2=$ Second year master degree, $3=$ Master of philosophy (mphil), $4=\mathrm{PhD}, 5=$ Other. Language: $1=$ English, $2=$ Other. Religion: $1=$ Muslim, $2=$ Hindu, $3=$ Sikh, $4=$ Christian, $5=$ Buddhist, $6=$ Parsi, $7=$ Other, $8=$ Prefer not say. Caste: $1=$ Scheduled caste, $2=$ Scheduled tribe, $3=$ Other backward castes, $4=$ General, $5=$ Other, $6=$ Prefer not say.

Table C.2: Summary statistics of variables in main econometric specification

|  | Mean | Standard Deviation |
| :--- | :---: | :---: |
| Outcome | 0.45 | 0.50 |
| Female Agent | 0.55 | 0.50 |
| Female Agent $\times$ High Outcome | 0.25 | 0.43 |
| Female Principal | 0.73 | 0.44 |
| Female Principal $\times$ High Outcome | 0.32 | 0.47 |
| Same Gender | 0.55 | 0.50 |
| N | 804 | 804 |

Notes: Table shows descriptive statistics of the corresponding variables.


Figure C.1: Mean payments to each party by realized outcome


Figure C.2: Mean payments to each party by realized outcome and gender of the agent

Table C.3: Regression results for principals' payments with session fixed effects

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $101.64^{* * *}$ | $102.27^{* * *}$ | $114.19^{* * *}$ | $112.27^{* * *}$ |
|  | $(14.80)$ | $(18.17)$ | $(33.43)$ | $(33.39)$ |
| Female Agent |  | 3.27 | 2.74 | 8.61 |
|  |  | $(9.08)$ | $(8.69)$ | $(9.30)$ |
| Female Agent $\times$ High Outcome |  | -1.09 | -0.17 | -0.96 |
|  |  | $(13.31)$ | $(12.43)$ | $(12.50)$ |
| Female Principal |  | -34.96 | -34.49 |  |
|  |  | $(36.85)$ | $(36.53)$ |  |
| Female Principal $\times$ High Outcome |  | -16.55 | -14.53 |  |
|  |  |  | $(35.06)$ | $(34.95)$ |
| Same Gender |  |  |  | $-11.79^{*}$ |
|  |  |  |  | $(6.45)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.37 | 0.37 | 0.39 | 0.39 |
| N | 804 | 804 | 804 | 804 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Session fixed effects are included in all specifications. Standard errors are clustered at the principal level.

Table C.4: Regression results for principals' payments with round fixed effects

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $101.25^{* * *}$ | $99.94^{* * *}$ | $111.83^{* * *}$ | $109.31^{* * *}$ |
|  | $(14.05)$ | $(17.39)$ | $(32.71)$ | $(32.76)$ |
| Female Agent |  | 0.92 | 0.31 | 8.64 |
|  |  | $(9.60)$ | $(9.40)$ | $(9.67)$ |
| Female Agent $\times$ High Outcome |  | 2.39 | 3.31 | 2.08 |
|  |  | $(13.45)$ | $(12.67)$ | $(12.85)$ |
| Female Principal |  | -33.23 | -32.75 |  |
|  |  |  | $(41.41)$ | $(40.53)$ |
| Female Principal $\times$ High Outcome |  | -16.17 | -13.29 |  |
|  |  |  | $(34.60)$ | $(34.57)$ |
| Same Gender |  |  | $-16.45^{* *}$ |  |
|  |  |  |  | $(6.86)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.34 | 0.34 | 0.35 | 0.35 |
| N | 804 | 804 | 804 | 804 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Round fixed effects are included in all specifications. Standard errors are clustered at the principal level.

Table C.5: Regression results for principals' payments without controls

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $108.21^{* * *}$ | $104.21^{* * *}$ | $116.58^{* * *}$ | $113.76^{* * *}$ |
|  | $(14.72)$ | $(18.76)$ | $(33.33)$ | $(33.61)$ |
| Female Agent |  | -4.81 | -1.34 | 7.22 |
|  |  | $(11.45)$ | $(11.79)$ | $(11.49)$ |
| Female Agent $\times$ High Outcome |  | 7.33 | 3.70 | 2.35 |
|  |  | $(16.98)$ | $(16.16)$ | $(16.28)$ |
| Female Principal |  | -27.71 | -29.09 |  |
|  |  | $(29.27)$ | $(29.01)$ |  |
| Female Principal $\times$ High Outcome |  | -16.22 | -12.93 |  |
|  |  | $(36.14)$ | $(36.23)$ |  |
| Same Gender |  |  | $-17.18^{* *}$ |  |
|  | 0.25 | 0.25 | 0.27 | 0.28 |
| R-Squared | 804 | 804 | 804 | 804 |
| N |  |  |  |  |

No controls are added to the regressions. Standard errors are clustered at the principal level.

Table C.6: Regression results for principals' payments for first ten rounds only

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $100.94^{* * *}$ | $101.99^{* * *}$ | $117.64^{* * *}$ | $111.16^{* * *}$ |
|  | $(15.54)$ | $(18.85)$ | $(38.08)$ | $(39.91)$ |
| Female Agent |  | 0.52 | -0.23 | 12.34 |
|  |  | $(15.60)$ | $(16.29)$ | $(20.80)$ |
| Female Agent $\times$ High Outcome |  | -1.93 | -1.19 | -1.13 |
|  |  | $(20.34)$ | $(20.77)$ | $(21.11)$ |
| Female Principal |  | -22.24 | -19.78 |  |
|  |  | $(47.32)$ | $(46.91)$ |  |
| Female Principal $\times$ High Outcome |  | -21.33 | -15.34 |  |
|  |  | $(41.91)$ | $(44.04)$ |  |
| Same Gender |  |  |  | $-24.09^{*}$ |
|  |  |  |  | $(14.24)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.31 | 0.31 | 0.32 | 0.33 |
| N | 420 | 420 | 420 | 420 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Results for only the initial 10 rounds are shown. Standard errors are clustered at the principal level.

Table C.7: Regression results for principals' payments after removing the first five rounds

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $109.94^{* * *}$ | $117.91^{* * *}$ | $129.97^{* * *}$ | $128.10^{* * *}$ |
| Female Agent | $(15.05)$ | $(18.59)$ | $(32.08)$ | $(31.87)$ |
|  |  | 7.11 | 5.49 | 15.25 |
| Female Agent $\times$ High Outcome |  | $(11.03)$ | $(10.72)$ | $(10.36)$ |
|  | -14.39 | -11.97 | -13.21 |  |
| Female Principal | $(14.31)$ | $(13.42)$ | $(13.51)$ |  |
|  |  |  | -36.68 | -35.99 |
| Female Principal $\times$ High Outcome |  |  | $(41.01)$ | $(39.37)$ |
|  |  |  | -17.68 | -15.89 |
| Same Gender |  |  |  | $(34.19)$ |
|  |  |  |  | $-20.11^{* *}$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $(7.41)$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.37 | 0.37 | 0.38 | 0.38 |
| N | 594 | 594 | 594 | 594 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. First five rounds were removed for regressions above. Standard errors are clustered at the principal level.

Table C.8: Regression results for principals' beliefs with session fixed effects

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.21^{* * *}$ | $0.22^{* * *}$ | $0.14^{* * *}$ | $0.14^{* * *}$ |
|  | $(0.02)$ | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Agent |  | 0.02 | 0.03 | $0.04^{* *}$ |
|  |  | $(0.02)$ | $(0.02)$ | $(0.02)$ |
| Female Agent $\times$ High Outcome |  | -0.02 | -0.02 | -0.03 |
|  |  | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Principal |  | $-0.14^{* *}$ | $-0.14^{* *}$ |  |
|  |  | $(0.05)$ | $(0.05)$ |  |
| Female Principal $\times$ High Outcome |  |  | $0.11^{* *}$ | $0.11^{* *}$ |
|  |  |  | $(0.04)$ | $(0.04)$ |
| Same Gender |  |  |  | -0.03 |
|  | $\checkmark$ | $\checkmark$ | $\checkmark$ | $(0.02)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | 0.33 | 0.34 | 0.34 | 0.35 |
| R-Squared | 804 | 804 | 804 | 804 |
| N |  |  |  |  |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Session fixed effects are included in all specifications. Standard errors are clustered at the principal level.

Table C.9: Regression results for principals' beliefs with round fixed effects

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.22^{* * *}$ | $0.23^{* * *}$ | $0.16^{* * *}$ | $0.16^{* * *}$ |
|  | $(0.03)$ | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Agent |  | 0.02 | 0.03 | $0.04^{*}$ |
|  |  | $(0.03)$ | $(0.03)$ | $(0.02)$ |
| Female Agent $\times$ High Outcome |  | -0.02 | -0.03 | -0.03 |
|  |  | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Principal |  | $-0.14^{* *}$ | $-0.14^{* *}$ |  |
|  |  |  | $(0.05)$ | $(0.05)$ |
| Female Principal $\times$ High Outcome |  |  | $0.10^{* *}$ | $0.10^{* *}$ |
|  |  |  | $(0.05)$ | $(0.05)$ |
| Same Gender |  |  |  | -0.02 |
|  | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | 0.33 | 0.34 | 0.35 | 0.35 |
| R-Squared | 804 | 804 | 804 | 804 |
| N |  |  |  |  |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Session fixed effects are included in all specifications. Standard errors are clustered at the principal level.

Table C.10: Regression results for principals' beliefs without controls

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.17^{* * *}$ | $0.18^{* * *}$ | $0.12^{* * *}$ | $0.12^{* * *}$ |
|  | $(0.03)$ | $(0.04)$ | $(0.04)$ | $(0.04)$ |
| Female Agent |  | 0.02 | 0.02 | 0.03 |
|  |  | $(0.02)$ | $(0.03)$ | $(0.03)$ |
| Female Agent $\times$ High Outcome |  | -0.02 | -0.02 | -0.02 |
|  |  | $(0.04)$ | $(0.04)$ | $(0.04)$ |
| Female Principal |  | -0.01 | -0.01 |  |
|  |  |  | $(0.06)$ | $(0.06)$ |
| Female Principal $\times$ High Outcome |  |  | $0.08^{*}$ | $0.08^{*}$ |
|  |  |  | $(0.05)$ | $(0.05)$ |
| Same Gender |  |  |  | -0.02 |
|  | $\checkmark$ | $\checkmark$ | $\checkmark$ | $(0.02)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | 0.25 | 0.25 | 0.25 | 0.25 |
| R-Squared | 804 | 804 | 804 | 804 |
| N |  |  |  |  |

No controls are included in the regressions. Standard errors are clustered at the principal level.

Table C.11: Regression results for principals' beliefs for the first ten rounds only

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.23^{* * *}$ | $0.25^{* * *}$ | $0.21^{* * *}$ | $0.20^{* * *}$ |
|  | $(0.04)$ | $(0.04)$ | $(0.04)$ | $(0.04)$ |
| Female Agent |  | 0.04 | 0.04 | 0.06 |
|  |  | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Agent $\times$ High Outcome |  | -0.03 | -0.04 | -0.04 |
|  |  | $(0.05)$ | $(0.05)$ | $(0.05)$ |
| Female Principal |  | $-0.15^{*}$ | $-0.15^{* *}$ |  |
|  |  | $(0.07)$ | $(0.07)$ |  |
| Female Principal $\times$ High Outcome |  |  | 0.06 | 0.07 |
|  |  |  | $(0.06)$ | $(0.06)$ |
| Same Gender |  |  |  | -0.03 |
|  | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | 0.31 | 0.31 | 0.33 | 0.33 |
| R-Squared | 420 | 420 | 420 | 420 |
| N |  |  |  |  |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Results for only the initial 10 rounds are shown. Standard errors are clustered at the principal level.

Table C.12: Regression results for principals' beliefs after removing the first five rounds

|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Outcome | $0.23^{* * *}$ | $0.22^{* * *}$ | $0.15^{* * *}$ | $0.15^{* * *}$ |
|  | $(0.02)$ | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Agent |  | 0.01 | 0.02 | 0.02 |
|  |  | $(0.02)$ | $(0.02)$ | $(0.02)$ |
| Female Agent $\times$ High Outcome |  | 0.00 | -0.00 | -0.00 |
|  |  | $(0.03)$ | $(0.03)$ | $(0.03)$ |
| Female Principal |  | $-0.14^{* *}$ | $-0.14^{* *}$ |  |
|  |  | $(0.05)$ | $(0.05)$ |  |
| Female Principal $\times$ High Outcome |  |  | 0.11 | $0.11^{* * *}$ |
|  |  |  | $(0.04)$ | $(0.04)$ |
| Same Gender |  |  |  | -0.02 |
|  | $\checkmark$ | $\checkmark$ | $\checkmark$ | $(0.02)$ |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | 0.37 | 0.36 | 0.38 | 0.38 |
| R-Squared | 594 | 594 | 594 | 594 |
| N |  |  |  |  |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Results above are shown after removing first five rounds of the sessions. Standard errors are clustered at the principal level.

Table C.13: Regression results for different parts of the cumulative distribution of payments

|  | $(\leq 50)$ | $(\leq 100)$ | $(\leq 150)$ | $(\leq 200)$ | $(\leq 250)$ | $(\leq 300)$ |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Outcome | $18.99^{* * *}$ | $21.53^{* * *}$ | $30.65^{* * *}$ | $43.74^{* * *}$ | $47.85^{* * *}$ | $67.66^{* * *}$ |
|  | $(5.59)$ | $(9.47)$ | $(11.49)$ | $(9.10)$ | $(8.79)$ | $(9.41)$ |
| Female Agent | -1.17 | -4.29 | -0.53 | 3.63 | -3.36 | 0.82 |
|  | $(3.31)$ | $(5.45)$ | $(6.83)$ | $(7.17)$ | $(7.81)$ | $(8.34)$ |
| Female Agent $\times$ | 2.73 | 6.94 | -0.13 | -4.94 | 5.59 | 10.79 |
| High Outcome | $(6.55)$ | $(11.08)$ | $(14.12)$ | $(11.72)$ | $(11.51)$ | $(12.27)$ |
|  |  |  |  |  |  |  |
| Mean payment | 17.00 | 43.79 | 70.21 | 118.99 | 139.84 | 161.94 |
| Demographics | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Task Controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| R-Squared | 0.58 | 0.49 | 0.42 | 0.34 | 0.34 | 0.33 |
| N | 150 | 224 | 300 | 510 | 608 | 706 |

Demographic variables include: principal's age, religion, caste, main language, state, education level, and field of study. Task controls are dummy variables for each task. Robust standard errors are reported.

## Bibliography

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011): "Reference points and effort provision," American Economic Review, 101, 470-92.

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.

Alan, S., S. Ertac, E. Kubilay, and G. Loranth (2020): "Understanding gender differences in leadership," The Economic Journal, 130, 263-289.

Alan, S., S. Ertac, and I. Mumcu (2018):"Gender stereotypes in the classroom and effects on achievement," Review of Economics and Statistics, 100, 876-890.

Ambuehl, S. and S. Li (2018): "Belief updating and the demand for information," Games and Economic Behavior, 109, 21-39.

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 $\mathcal{E}^{3}$ 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.

Azmat, G. and N. Iriberri (2010): "The importance of relative performance feedback information: Evidence from a natural experiment using high school students," Journal of Public Economics, 94, 435-452.

Azmat, G. and B. Petrongolo (2014): "Gender and the labor market: What have we learned from field and lab experiments?" Labour Economics, 30, 32-40.

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.

Banerjee, R., N. D. Gupta, and M. C. Villeval (2020): "Feedback spillovers across tasks, self-confidence and competitiveness," Games and Economic Behavior, 123, 127-170.

Barber, B. M. and T. Odean (2001): "Boys will be boys: Gender, overconfidence, and common stock investment," The quarterly journal of economics, 116, 261-292.

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.

Benjamin, D. (2018): "Errors in Probabilistic Reasoning and Judgment Biases," Working Paper.

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.

Blanes i Vidal, J. and M. Nossol (2011):"Tournaments without prizes: Evidence from personnel records," Management Science, 57, 1721-1736.

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.

Block, R. A. and D. R. Harper (1991): "Overconfidence in estimation: Testing the anchoring-and-adjustment hypothesis," Organizational Behavior and Human Decision Processes, 49, 188-207.

Bohnet, I., A. Van Geen, and M. Bazerman (2016): "When performance trumps gender bias: Joint vs. separate evaluation," Management Science, 62, 1225-1234.

Bohren, J. A., A. Imas, and M. Rosenberg (2019): "The dynamics of discrimination: Theory and evidence," American Economic Review, 109, 3395-3436.

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 (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 D. Y. Yang (2021):"Misperceptions about Others,".
Buser, T., M. Niederle, and H. Oosterbeek (2014): "Gender, competitiveness, and career choices," The quarterly journal of economics, 129, 1409-1447.

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 28233338.

Carlana, M. (2019): "Implicit stereotypes: Evidence from teachers' gender bias," The Quarterly Journal of Economics, 134, 1163-1224.

Castagnetti, A. and R. Schmacker (2020): "Protecting the Ego: Motivated Information Selection," Working Paper.

Charness, G. and U. Gneezy (2012): "Strong evidence for gender differences in risk taking," Journal of Economic Behavior E Organization, 83, 50-58.

Charness, G., R. Oprea, and S. Yuksel (2018): "How Do People Choose Between Biased Information Sources? Evidence from a Laboratory Experiment," Working Paper.

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

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree - An open-source platform for laboratory, online, and field experiments," Journal of Behavioral and Experimental Finance, 9, 88-97.

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 (2019a): "Stereotypes and Belief Updating," Working Paper.
_ (2019b): "When to Apply?" Working Paper, Harvard Business School.
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.

Croson, R. and U. Gneezy (2009): "Gender differences in preferences," Journal of Economic Literature, 47, 448-74.

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,"

Danz, D., A. J. Wilson, and L. Vesterlund (2020): "Belief Elicitation: Limiting Truth Telling with Information or Incentives," CESifo Working Paper.

Dean, J. T. and S. Jayachandran (2019): "Changing Family Attitudes to Promote Female Employment," in AEA Papers and Proceedings, vol. 109, 13842.

Deaux, K. and T. Emswiller (1974):"Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female." Journal of Personality and Social Psychology, 29, 80-85.

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.

Dobrescu, L. I., M. Faravelli, R. Megalokonomou, and A. Motta (2021): "Relative performance feedback in education: Evidence from a randomised controlled trial," The Economic Journal, 131, 3145-3181.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," Journal of the European Economic Association, 9, 522550.

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.

Dweck, C. S., W. Davidson, S. Nelson, and B. Enna (1978): "Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis." Developmental Psychology, 14, 268-276.

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

Eil, D. and J. M. Rao (2011): "The good news-bad news effect: asymmetric processing of objective information about yourself," American Economic Journal: Microeconomics, 3, 114-38.

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.

Espinoza, P., A. B. A. da Luz Fontes, and C. J. Arms-Chavez (2014): "Attributional gender bias: Teachers' ability and effort explanations for students' math performance," Social Psychology of Education, 17, 105-126.

Exley, C. L. and J. B. Kessler (2018):"Motivated Errors," Tech. rep.
— (2019): "The gender gap in self-promotion," Tech. rep., National Bureau of Economic Research.

Falk, A., F. Kosse, H. Schildberg-Hörisch, and F. Zimmermann (2020): "Self-assessment: The role of the social environment," Working Paper.

Falk, A. and F. Zimmermann (2016): "Beliefs and utility: Experimental evidence on preferences for information," IZA Discussion Paper No. 10172.

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.

Flory, J. A., A. Leibbrandt, and J. A. List (2015): "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions," The Review of Economic Studies, 82, 122-155.

Folke, O. and J. K. Rickne (2020): "Sexual Harassment and Gender Inequality in the Labor Market," CEPR Discussion Paper No. DP1473\%.

Ganguly, A. and J. Tasoff (2016): "Fantasy and dread: the demand for information and the consumption utility of the future," Management Science, 63, 4037-4060.

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

Gneezy, U. and J. Potters (1997): "An experiment on risk taking and evaluation periods," The Quarterly Journal of Economics, 112, 631-645.

Goldin, C. (2014): "A grand gender convergence: Its last chapter," American Economic Review, 104, 1091-1119.

Golman, R., D. Hagmann, and G. Loewenstein (2017): "Information avoidance," Journal of Economic Literature, 55, 96-135.

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.

Gurdal, M. Y., J. B. Miller, and A. Rustichini (2013): "Why blame?" Journal of Political Economy, 121, 1205-1247.

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. $2562 \%$.

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 SelfEmployed 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," .

Jones, E. E. and V. A. Harris (1967):"The attribution of attitudes," Journal of Experimental Social Psychology, 3, 1-24.

Jones, T. M., K. V. Fanson, R. Lanfear, M. R. Symonds, and M. Higgie (2014): "Gender differences in conference presentations: A consequence of selfselection?" PeerJ, 2, e627.

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

Karlsson, N., G. Loewenstein, and D. Seppi (2009): "The ostrich effect: Selective attention to information," Journal of Risk and Uncertainty, 38, 95-115.

Karni, E. (2009): "A mechanism for eliciting probabilities," Econometrica, 77, 603-606.

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 ViolenceEvidence from a Field Experiment in Ethiopia," .

Landsman, R. (2019): "Topics in Labor and Experimental Economics," Ph.D. thesis, University of Pittsburgh.

Lindo, J. M., P. Siminski, and I. D. Swensen (2018): "College Party Culture and Sexual Assault," American Economic Journal: Applied Economics, 10, 23665.

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.

Masatlioglu, Y., A. Y. Orhun, and C. Raymond (2017):"Intrinsic information preferences and skewness," Working Paper.

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," .

Michelmore, K. and S. Sassler (2016): "Explaining the gender wage gap in STEM: Does field sex composition matter?" RSF: The Russell Sage Foundation Journal of the Social Sciences, 2, 194-215.

Milkman, K., M. Akinola, and D. Chugh (2013): "Discrimination in the Academy: A Field Experiment," SSRN, Working Paper.

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

Mobius, M., M. Niederle, P. Niehaus, and T. Rosenblat (2014): "Managing self-confidence: Theory and experimental evidence," Working Paper.

Montanari, G. and S. Nunnari (2019): "Audi Alteram Partem: An Experiment on Selective Exposure to Information," Working Paper.

Moore, D. A. and P. J. Healy (2008): "The trouble with overconfidence," Psychological Review, 115, 502.

Moore, D. A., S. A. Swift, Z. S. Sharek, and F. Gino (2010): "Correspondence bias in performance evaluation: Why grade inflation works," Personality and Social Psychology Bulletin, 36, 843-852.

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

Niederle, M. and L. Vesterlund (2007): "Do women shy away from competition? Do men compete too much?" The Quarterly Journal of Economics, 122, 1067-1101.

Nielsen, K. (2018): "Preferences for the resolution of uncertainty and the timing of information," Working Paper.

Oster, E., I. Shoulson, and E. Dorsey (2013): "Limited life expectancy, human capital and health investments," American Economic Review, 103, 19772002.

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.

Ross, L. (1977): "The intuitive psychologist and his shortcomings: Distortions in the attribution process," in Advances in experimental social psychology, Elsevier, vol. 10, 173-220.

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.

Sarsons, H. (2017): "Recognition for group work: Gender differences in academia," American Economic Review, 107, 141-45.

- (2019): "Interpreting signals in the labor market: evidence from medical referrals," Working Paper.

Selody, K. (2010): "Board independence and the gender pay gap for top executives," Working Paper.

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2015): "Financial attention," The Review of Financial Studies, 29, 863-897.

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.

Zimmermann, F. (2014): "Clumped or piecewise? Evidence on preferences for information," Management Science, 61, 740-753.

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


[^0]:    ${ }^{1}$ 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.

[^1]:    ${ }^{7}$ Reporting of sexual harassment can lead to victim blaming, backlash from the perpetrator and other forms of retaliation (Dahl and Knepper (2021)).

[^2]:    ${ }^{8}$ 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).

[^3]:    ${ }^{10}$ 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.

[^4]:    ${ }^{11}$ See Figure A. 1 in the appendix
    ${ }^{12} \mathrm{https}: / /$ www.rainn.org/statistics/campus-sexual-violence
    ${ }^{13}$ Elaborated in appendix A. 1

[^5]:    ${ }^{14}$ 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.

[^6]:    ${ }^{16}$ 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.

[^7]:    ${ }^{17}$ Feedback from trainers after the training revealed that men liked the candid nature of this training.
    ${ }^{18}$ 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.

[^8]:    ${ }^{19}$ 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.
    ${ }^{20}$ This was a subset of the intervention undertaken for men and provided women with information on how to detect sexual harassment.

[^9]:    ${ }^{21}$ This is because the female intervention college came on-board for the project much later than the two colleges for male intervention.

[^10]:    ${ }^{22}$ 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.

[^11]:    ${ }^{24}$ 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 1.5, I do not find any difference in awareness about sexual harassment between women in treatment and control classes
    ${ }^{26}$ Female surveyors help in making women participants more comfortable when answering sen-

[^12]:    sitive 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.
    ${ }^{30}$ Same gender groups also help with obfuscation (Haaland et al. (2020)), reducing the threat of demand effects.

[^13]:    ${ }^{31}$ 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.

[^14]:    ${ }^{32}$ 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.

[^15]:    ${ }^{34}$ I show in the appendix table A. 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).

[^16]:    ${ }^{37}$ 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

[^17]:    impute these values (Kling et al. (2007))

[^18]:    ${ }^{38}$ Although women too become more aware about ICC in training classes by $6.3 \mathrm{p} . \mathrm{p}$ (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.
    ${ }^{39}$ 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.

[^19]:    ${ }^{40}$ 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.

[^20]:    ${ }^{1}$ Academic settings or modern workplaces are some common avenues where public gatherings (e.g. presentations) are useful not only to receive performance feedback but also for finding potential collaborations and partnerships.
    ${ }^{2}$ This is discussed in detail in section 2.3.3

[^21]:    ${ }^{3}$ See Golman et al. (2017) for a review of this literature.
    ${ }^{4}$ Similarly, we contribute to the literature that studies how individuals process the signals received from exogenously given information structures. Specifically, this literature studies whether there are asymmetries in the way individuals process positive and negative information about their performance (i.e. the literature on asymmetric updating). See Benjamin (2018) for a review.

[^22]:    ${ }^{5}$ For examples in the applied literature on information avoidance in financial settings see Karlsson et al. (2009) and Sicherman et al. (2015); for examples in medical contexts see Ganguly and Tasoff (2016) and Oster et al. (2013).
    ${ }^{6}$ Moreover, in T1, T2 and T3, we shut down any possibility of partners affecting one's payoffs
    ${ }^{7}$ For a comprehensive review see Croson and Gneezy (2009).

[^23]:    ${ }^{8}$ Screenshot of the experiment are provided in Appendix B.1.1.
    ${ }^{9}$ In case of ties, they were broken randomly. Subjects were informed about this feature.
    ${ }^{10}$ For a detailed description of this method, we refer the interested reader to Karni (2009).

[^24]:    ${ }^{11}$ However, if any subject demanded more information about the method, they could click on a button to read a full description of the elicitation method.

[^25]:    ${ }^{12}$ Remember that subjects were randomly matched in pairs of two after the prior belief elicitation.

[^26]:    ${ }^{13}$ Again taking the example of academia, it is generally considered costly if researchers do not present their research to the scientific community in terms of future collaborations or job opportunities.

[^27]:    ${ }^{14}$ See Moore and Healy (2008) for a detailed description of the different ways in which individuals can be overconfident.

[^28]:    ${ }^{15}$ Higher the number, higher is the willingness to take risks

[^29]:    ${ }^{16}$ Throughout this section we conduct probit regressions along with a linear probability regression model since our outcome variable is dichotomous.

[^30]:    ${ }^{17}$ We do this since there are no significant differences across them (neither overall nor by gender), as shown in the previous analysis.

[^31]:    ${ }^{18}$ In Columns (1)-(3) we report the results of OLS regressions, while in columns (4)-(6) we report those for probit regressions. In Columns (2) and (5) we add subjects' prior beliefs as control variable, while in Columns (3) and (6) we add, on top, the demographic control variables.

[^32]:    ${ }^{19}$ Recall that these estimates keep the values of the controls at their mean levels and hence the magnitudes are larger than OLS.

[^33]:    ${ }^{20}$ Remember that it was 42.826 for the prior beliefs.

[^34]:    ${ }^{21}$ For below median performers, efficiency will be determined by a trade-off between learning about one's own true rank and the probability of being hired by the partner.

[^35]:    ${ }^{22}$ For instance when answers were elicited on scale of $[-1,1]$ where -1 is 'Women know more', 1 is 'Men know more' and 0 is labelled as ' no gender difference', men's average for art is -0.317 and for women is -0.419 . For sports however, men's average is 0.643 and for women is 0.571 .

[^36]:    ${ }^{1}$ Refer to Azmat and Petrongolo (2014) for a review of the literature.

[^37]:    ${ }^{2}$ That is, at every round principals and agents were randomly rematched.
    ${ }^{3}$ In two sessions, participants played 9 rounds per task, instead.

[^38]:    ${ }^{4}$ We opted for the stranger-matching protocol since it avoids reputation building and related strategic concerns.
    ${ }^{5}$ At the time of the experiment, this amount corresponded to $£ 3.92$ (exchange rate as of July 2018: $£ 1.00=₹ 89.21$ ).
    ${ }^{6}$ We implemented this payment procedure following the same considerations as in Gurdal et al. (2013). In particular, two features are worth noting. First, not allowing the principal to keep any unassigned money for herself shuts down any (financial) incentive for the principal to keep all the money. Second, having the option to also pay a random agent allow us to eliminate any efficiency motives (in terms of subjects versus experimenter considerations) that the principal might have and we can more tightly learn whether the principal is holding responsible that specific agent depending on the realized outcome.

[^39]:    ${ }^{7}$ Showing other religious groups or full names would have primed religion and/or caste.

[^40]:    ${ }^{8}$ The low output was always set equal to ₹ 0 .
    ${ }^{9}$ However, it could never reach $100 \%$ probability.

[^41]:    ${ }^{10}$ There are also no significant differences in performances' distributions nor in the variance of the number of correct questions by gender and across tasks.

[^42]:    ${ }^{11}$ Given that the principals' payments were made from a separate pot of money, and thus they were payoff irrelevant for the principals, one possibility could have been that these payments did not vary or were chosen randomly. However, the results of this section clearly shows that this was not the case: principals understood that their choices had economic implications for their matched agents and they took this payment decision seriously.
    ${ }^{12}$ The principals' beliefs correspond to the proportion of correct questions that they think the agents have solved correctly in one given round.

[^43]:    ${ }^{13}$ In the appendix, we also show mean payments (from the separate pot) made to the other randomly matched agent and to the experimenter by realized outcome and the gender of the agent.

[^44]:    ${ }^{14}$ The number of observations is 804 because in two sessions we had 9 rounds per task and hence 18 rounds instead of 20 . This gives us $180,160,140,162$ and 162 observations for each session.

[^45]:    ${ }^{15}$ In particular, we asked the following open-text question: "Please, guess what our research questions are."

[^46]:    ${ }^{16}$ Importantly, our subject pool was new to experiments. Therefore, they were not aware that in standard experiments subjects' personal characteristics (such as (nick-)names and age) are not usually disclosed.

[^47]:    ${ }^{1}$ It was not feasible to have third part observers since that would change behavior of the students even more

[^48]:    Once you click next, you will have 10 minutes to solve the Quiz.

