

# A man's world?

## The impact of a male dominated environment on female leadership



Andreas Born (Stockholm School of Economics)

Eva Ranehill (University of Gothenburg)

Anna Sandberg (Stockholm University)

Gender Economics Workshop, Berlin, October 19<sup>th</sup> 2018

# MOTIVATION

---

- Despite important advancement in women's labor market participation and educational attainment, labor markets remain highly segregated by gender
  - Vertically: Women are underrepresented at leading positions. (*e.g. Bertrand & Hallock 2001, Blau and Kahn 2017, Lawless and Fox 2012*)
  - Horizontally: Women are underrepresented in high-paying occupations (*e.g. Bettio & Verashchagina 2009, Cohen 2013, Olivetti & Petrongolo 2016*)

# MOTIVATION

---

- As a result, conventional human capital today explains only a small part of gender wage gap while occupation and industry remain important. (*Blau and Kahn 2016*)
- **Why do so few women reach the top, and instead persist in educations with less career potential, in lower paid industries and occupations, and at lower paid levels within firms?**
- Previous research has proposed different explanations such as discrimination, gender preference gaps, gender gaps in parental investment etc.

# MOTIVATION

---

- **This study: What if minority status, in itself, has an adverse impact on women's careers?**
- If so, the absence of women in male-dominated high status areas may be a **self-reinforcing process**:
  - Women become reluctant to enter, and prone to leave, such environments...
- If true this may help explain many prominent gender gaps in labor markets (e.g. vertical and horizontal gaps, gender tipping points, low female job satisfaction in male majority workplaces)

*(e.g. Goldin 2014, Blau & Kahn 2000, 2006, e.g. Pan 2015, Lordan & Pischke 2016, Hunt 2016)*

# OUR MAIN QUESTION

---

- This study focuses on gender gaps in willingness to become the leader of a team. Our primary research question is:

**Is women's willingness to lead negatively influenced by environments where men are in the majority?**

# PREVIOUS RESEARCH

---

- Women are (often) less likely than men to seek out leadership positions.

*(e.g. Preece and Stoddard 2015, Lawless & Fox 2011, Powell & Butterfield 2003, Kanthak and Woon 2015, Reuben et al. 2012, Ertac & Gurdal 2011, ...)*

- Environments with a large proportion of men may have an adverse effect on female performance, well-being and behavioral strategies...

*(e.g. Kanter 1977, Mendelberg & Karpowicz 2016, Eagly et al. 1992, Booth & Nolen 2012, Booth et al. 2014, Gneezy et al. 2003, Kubnen and Tymula 2012, Babcock et al. 2017, Hogarth et al. 2012, Bursztyn et al. 2017...)*

# OUR STUDY

---

## RESEARCH QUESTION

Is women's willingness to lead adversely affected by being in numerical minority in a team?

## METHOD

Laboratory experiment with exogenous variation in gender composition of teams where team members volunteer for leadership and vote for their preferred leader.

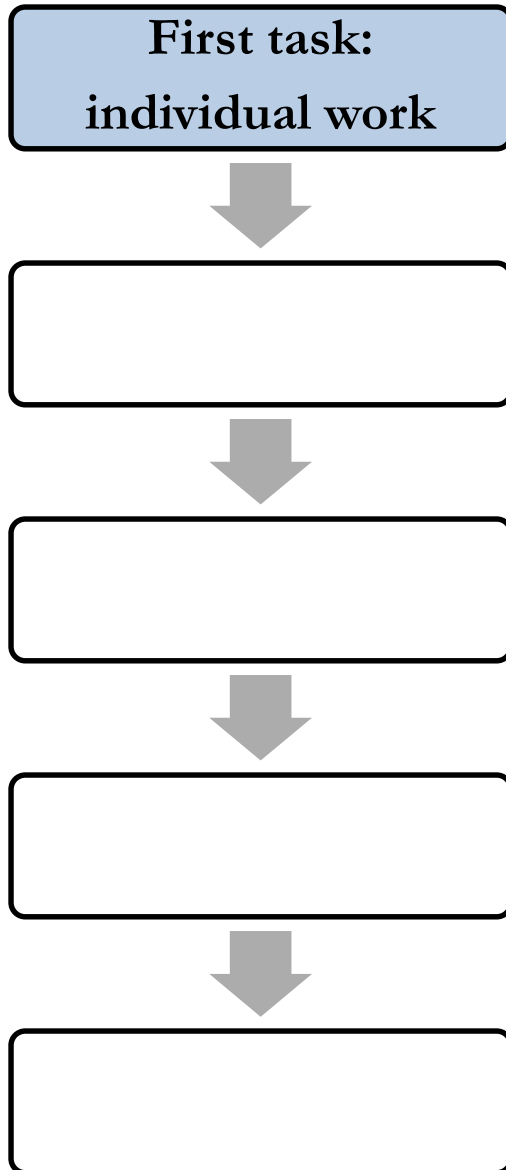
## PRIMARY HYPOTHESES

**H1** Women are more willing to become the leader of female majority teams than of male majority teams.

**H2** The gender difference in willingness to become the leader is smaller in female majority teams than in male majority teams.

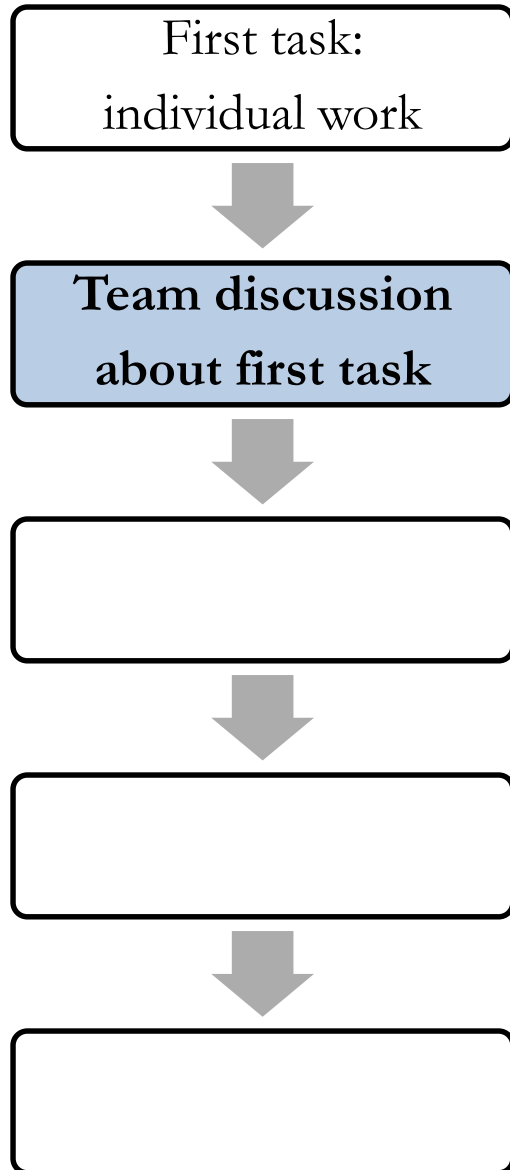
# EXPERIMENTAL DESIGN

---



- Participants work on a task individually on their computer.
- Task: Hypothetical scenario (“Lost at Sea”)
  - Alone on a life raft in the middle of the Atlantic after a boat accident.
  - **Rank 10 different items** in terms of their importance for survival.
- The participant’s solution is compared to that of a group of experts.
- **Closer to expert solution → higher payment.**

# EXPERIMENTAL DESIGN



- Participants are assigned to teams of four.
- Randomly assigned to one of two treatments:

**Male majority team**  
(3 men, 1 woman)

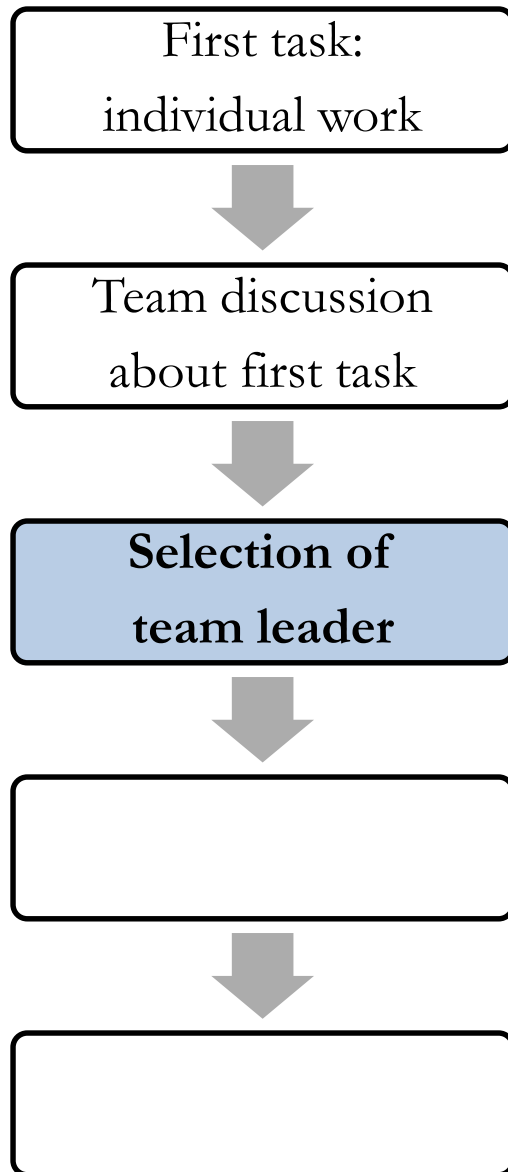


**Female majority team**  
(3 women, 1 man)

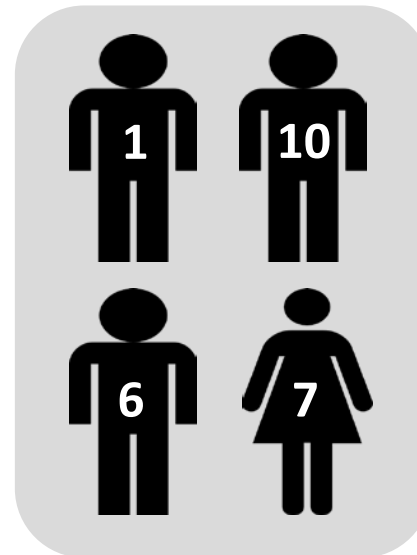


- Each team goes to a separate room and discusses the first task and comes up with a joint team solution.

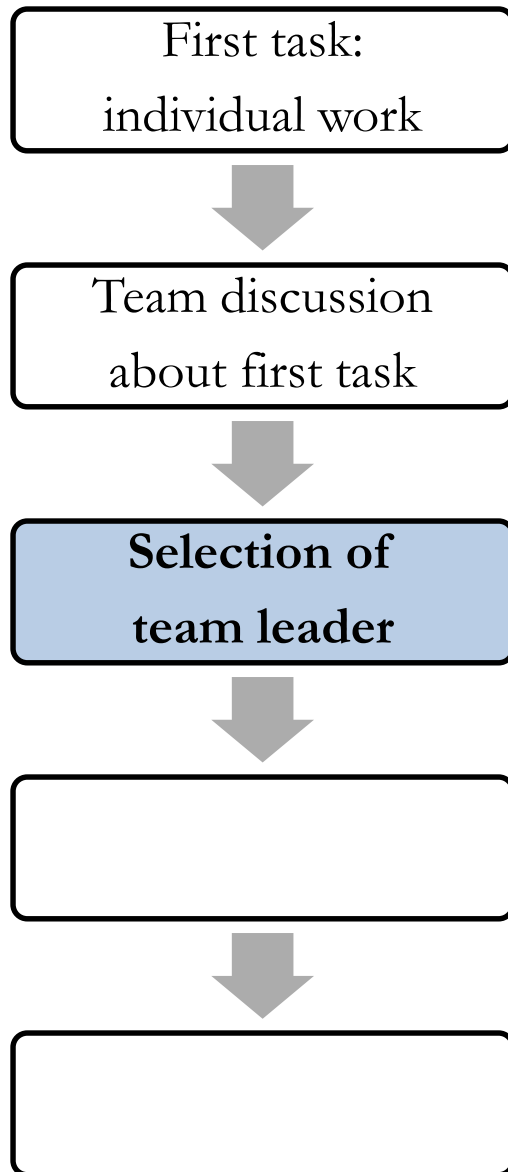
# EXPERIMENTAL DESIGN



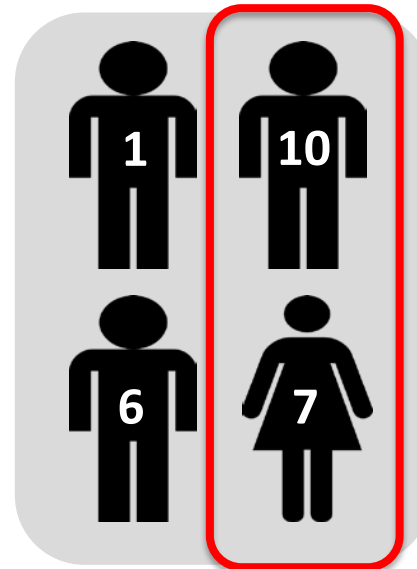
- Back at their computer: Informed that they must select a team leader before solving a second task.
- The job of the leader will be to decide on a joint team answer for the second task.
- **Each team member indicates how willing they are to become the team leader (1=not at all, 10=very much).**



# EXPERIMENTAL DESIGN

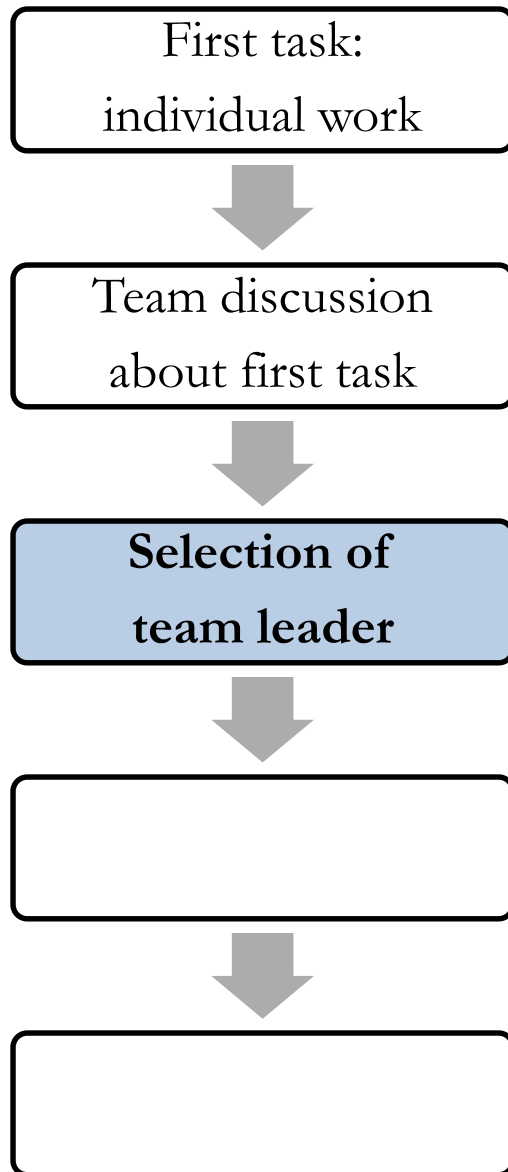


- The two team members who want to become leader the most become candidates in an election.
  - Randomized in case of ties
- **Stated willingness to lead has a direct impact of the likelihood to become the team leader**



# EXPERIMENTAL DESIGN

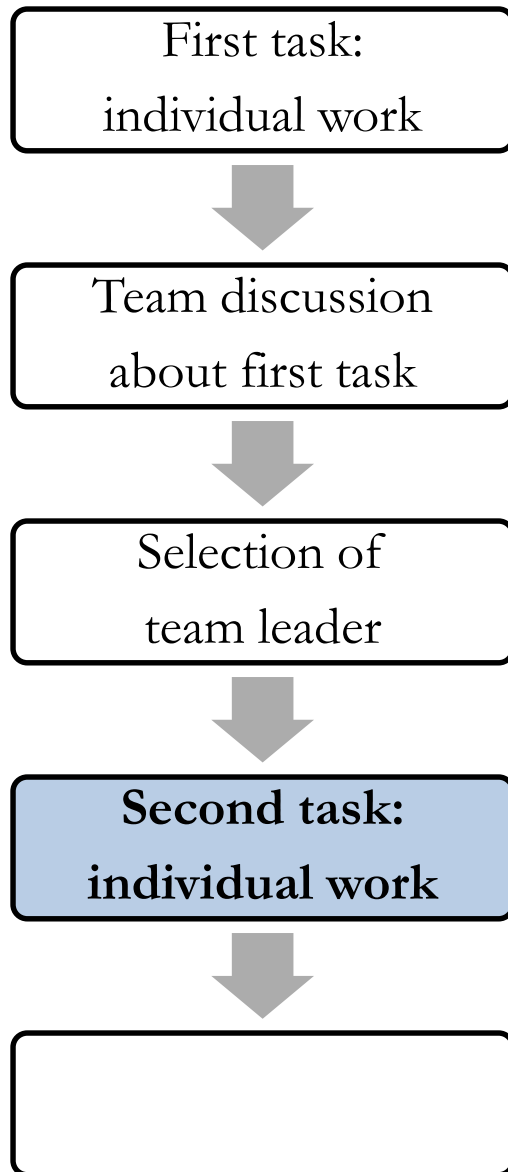
---



- Before it is revealed who the two candidates are, all team members vote.
  - Rank the other three team members in terms of who they would like to see as the leader.
- The votes provided by the two team members who are not candidates then determine which candidate becomes the team leader.

# EXPERIMENTAL DESIGN

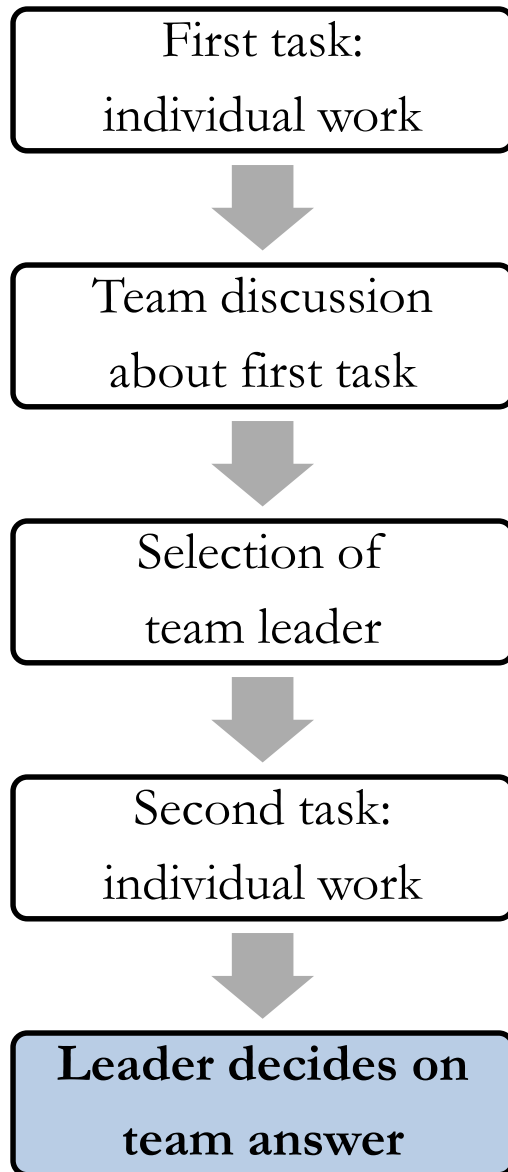
---



- Participants work individually on their computer with the second task.
- The second task is very similar to the first task, but with a different survival scenario (“Desert survival”).

# EXPERIMENTAL DESIGN

---



## Leader's task:

- The leader sees the suggested answers from all other team members.
- Based on this input, **the leader decides on a joint team answer.**

# EXPERIMENT DETAILS

---

- Participants are students from the University of Zurich and the ETH.
- In total 580 participants, divided into 145 teams.
- One randomly chosen task counts for payment.
- Pre-analysis plan published at the Open Science Framework.

---

	<b>Female majority group</b>	<b>Male majority group</b>	<b>Total</b>
<b>Women</b>	207	76	283
<b>Men</b>	69	228	297
<b>Total</b>	276	304	580

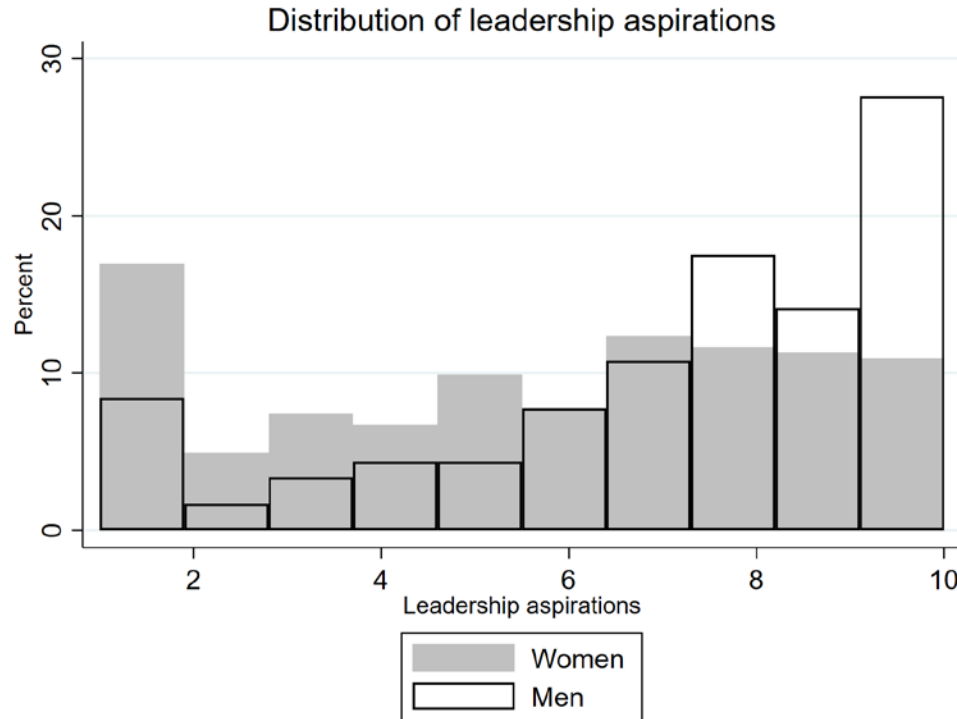
---

# RESULTS

---

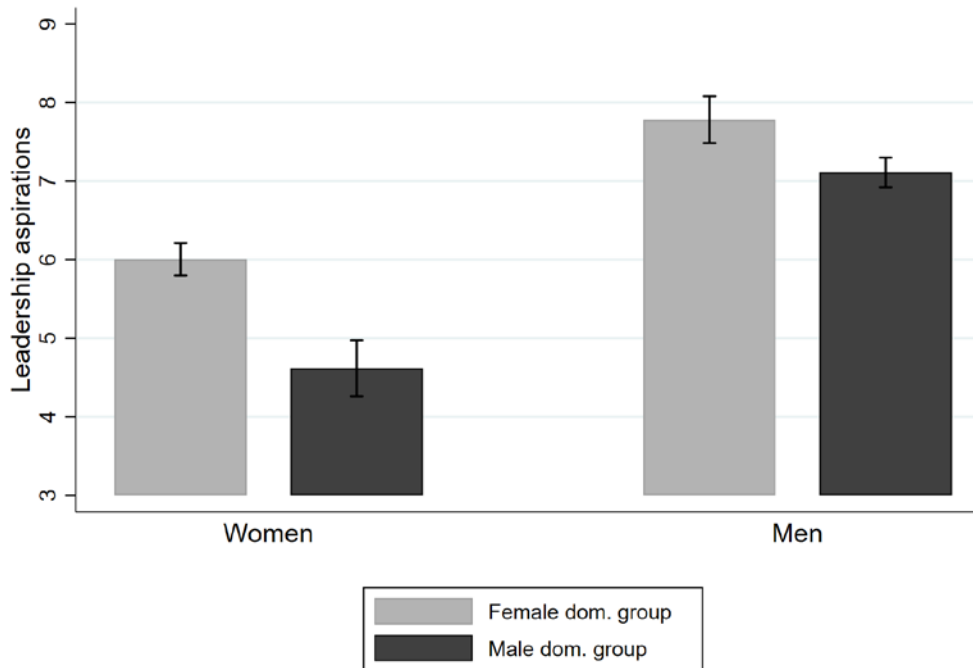
1. Baseline gender gap in willingness to lead
2. Main results: Impact of group gender composition
3. Mechanisms: Task related confidence, electoral outcomes, influence...
4. Outcomes

# GENDER GAP IN LEADERSHIP MOTIVATION



- On average, men are 1.63 units ( $d=0.56$ ) more willing than women to become the team leader ( $p<0.001$ ).
- The modal response for women is 1/10 (17%).
- The modal response for men is 10/10 (27%).
- KS-test:  $p<0.001$ .

# IMPACT OF GROUP GENDER COMPOSITION



Note: The error bars represent +/- 1 standard errors

- **H1:** Women are significantly more willing to become the leader in female majority teams than in male majority teams (increase from 4.62 to 6.00,  $p=0.001$ ,  $d=0.46$ ).
- Men are also more willing to become the leader in female majority teams, but not significant at the five percent level (increase from 7.11 to 7.78,  $p=0.133$ ).
- **H2:** The increase in leadership motivation is 0.71 units larger for women than for men, but this difference is not significant ( $p=0.231$ ).

# MAIN RESULTS

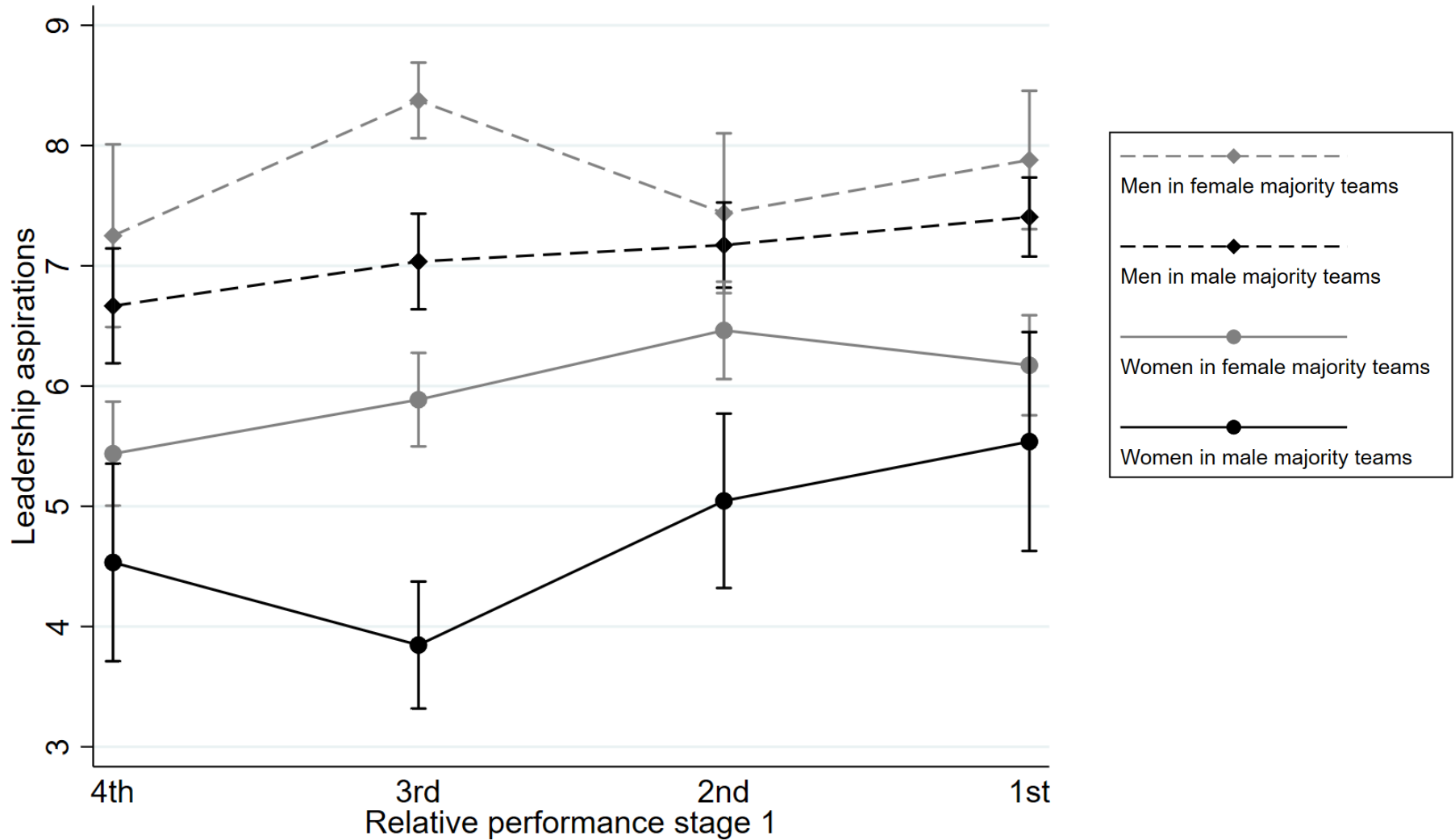
---

- On average, men perform slightly better than women at the "Lost at Sea" task.
- Are the gender and treatment differences in willingness to lead driven by gender differences in task performance?
  - NO
  - Controlling for relative performance, the gender and treatment effects remain stable.

[Link: Gender difference in performance](#)

[Link: Regression results](#)

# LEADERSHIP MOTIVATION BY GENDER, TREATMENT AND RELATIVE PERFORMANCE



# POTENTIAL MECHANISMS

---

- Why are women less willing to become the team leader in male majority teams than in female majority teams?
- We explore the following mechanisms
  - ✓ Confidence in own task performance
  - ✓ Support from team members
  - ✓ Anticipated support from team members
  - ✓ Influence in team discussions
  - ✓ Performance in the second task
  - ✓ Gender stereotypes held by team members
- First, we ask if these measures vary by gender and treatment.
- Then, we see how much the gender gap and treatment effect decline when controlling for each measure.

# CONFIDENCE

---

**Are women in male majority teams less confident in their relative ability than women in female majority teams?**

**Two measures of confidence:**

## **1. Guess of relative performance**

- After the group discussion all participants guess how well they did in the first task relative to the other group members (1st/2nd/3rd/4th).
- Paid a bonus of CHF 2 for correct guess.

# CONFIDENCE

---

**Are women in male majority teams less confident in their relative ability than women in female majority teams?**

## **2. Updating of individual answer after group discussion**

- After the group discussion all participants get the opportunity to update their previous answers in the first task.
- We compute each participant's degree of updating relative to the other team members.
  - (i) Sum (over all 10 items) the absolute difference in rank between the participant's initial answer and their updated answer.
  - (ii) Divide by the sum of all team member's differences.
- 0=the participant did not update at all.  
1=the participant was the only group member who updated.

# CONFIDENCE

---

## GENDER GAP

On average, controlling for relative performance, compared to men ...

... women **believe that their performance ranks 0.44 positions lower** ( $p < 0.001$ ).

... women are **more prone to update their answer** ( $p < 0.001$ ).

[Link: Regressions](#)

## TREATMENT EFFECT ON WOMEN

On average, controlling for relative performance, compared to women in female majority teams ...

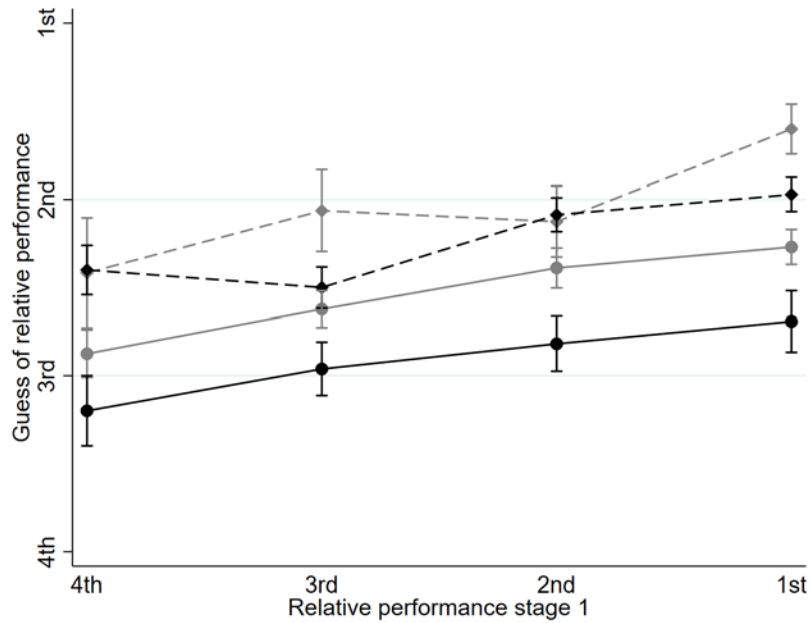
... women in male majority teams **believe that their performance ranks 0.37 positions lower** ( $p < 0.001$ ).

... women in male majority teams are **more prone to update their answer** ( $p = 0.006$ ).

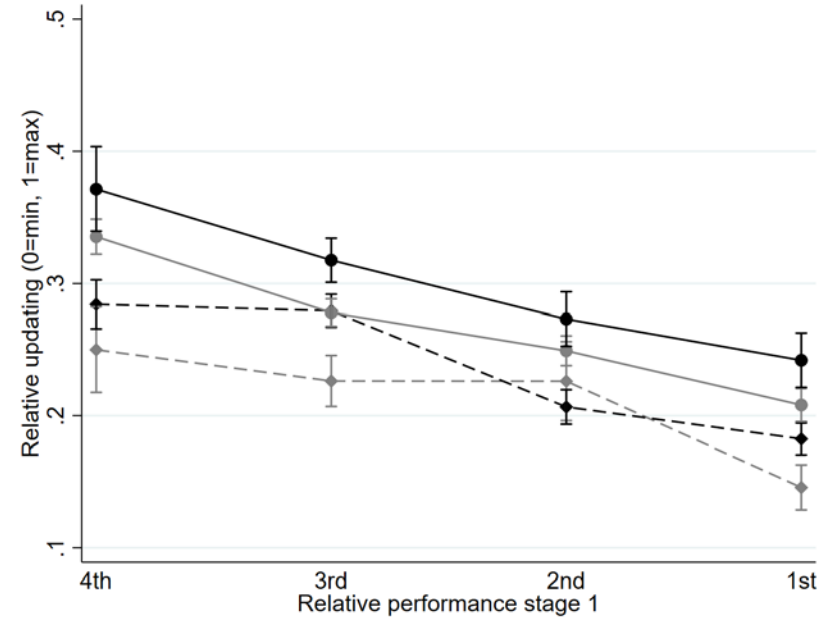
[Link: Regressions](#)

# CONFIDENCE

## GUESS OF OWN RELATIVE PERFORMANCE



## RELATIVE UPDATING



Men in female majority teams

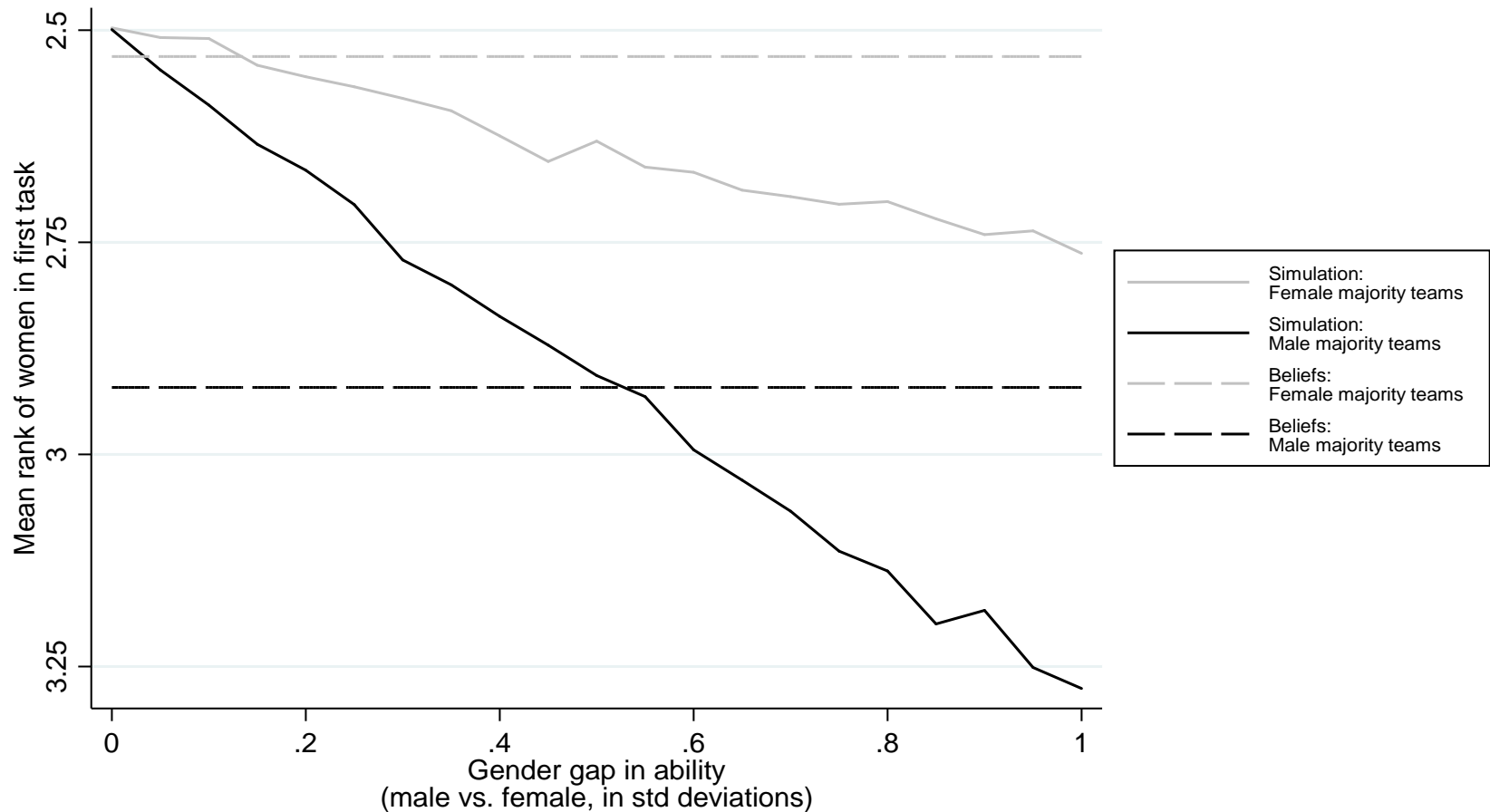
Men in male majority teams

Women in female majority teams

Women in male majority teams

# CONFIDENCE

Can relative performance beliefs account for the treatment effect in willingness to lead?



Simulations with 10,000 replications.

# ACTUAL AND ANTICIPATED SUPPORT

---

**Do women in male majority teams receive less support from their team members? Do they anticipate to receive less support?**

## **Rank in election**

- 1st/2nd/3rd/4th
- For this analysis: assigned average rank in case of ties.

## **Guess of rank in election**

- Before the results of the election are revealed, each participant must guess how well they will do in the election relative to the other group members (1st/2nd/3rd/4th).
- Paid a bonus of CHF 2 for correct guess.

# ACTUAL AND ANTICIPATED SUPPORT

---

## GENDER GAP

On average, controlling for relative performance, compared to men ...

... women are **ranked 0.42 positions worse** in the election ( $p < 0.001$ ).

... women *believe* that they will be ranked **0.58 positions worse** in the election ( $p < 0.001$ ).

[Link: Regressions](#)

## TREATMENT EFFECT ON WOMEN

On average, controlling for relative performance, compared to women in female majority teams ...

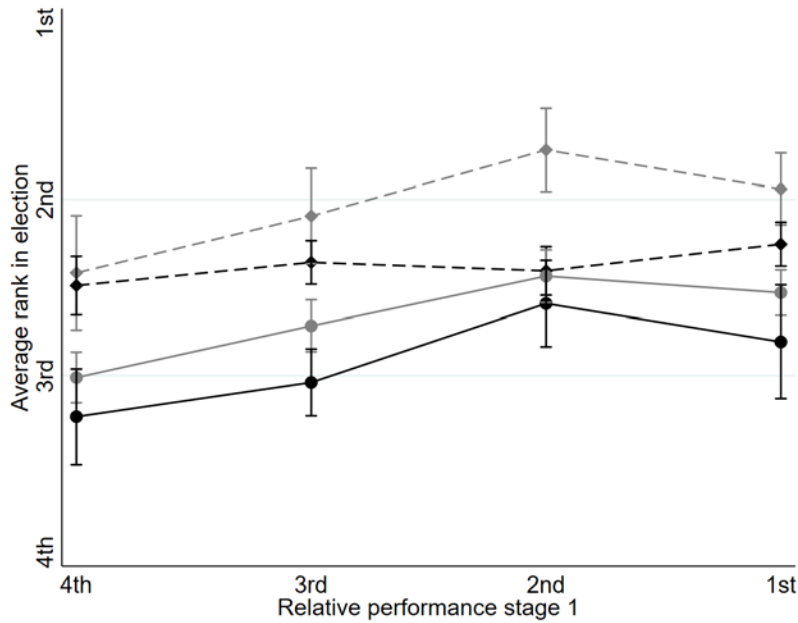
... women in male majority teams are ranked 0.23 positions worse in the election (but not significantly so,  $p = 0.080$ ).

... women in male majority teams *believe* that they will be ranked **0.26 positions worse** in the election ( $p = 0.020$ ).

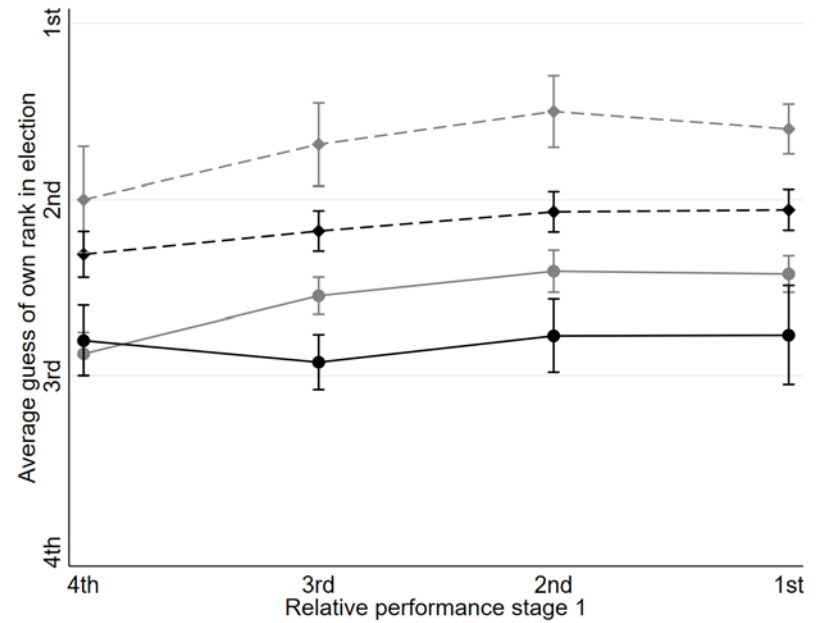
[Link: Regressions](#)

# ACTUAL AND ANTICIPATED SUPPORT

## ACTUAL RANK IN ELECTION



## GUESS OF RANK IN ELECTION



Men in female majority teams

Men in male majority teams

Women in female majority teams

Women in male majority teams

# INFLUENCE

---

Are women less influential in male majority teams than in female majority teams?

## Measure of influence

- How close is the team's joint answer to the participant's initial individual answer?
  - (i) Sum (over all ten items) the absolute difference in rank between the participant's individual answer and the team's joint answer.
  - (ii) Divide by the sum of all team members' differences.
  - (iii) Reverse the scale.
- **0= No influence:**

The team's answer is different from the participant's initial answer but identical to that of the other team members.
- **1= Maximum influence:**

The team's solution is identical to the participant's initial answer.

# INFLUENCE

---

## GENDER GAP

On average, controlling for relative performance, compared to men ...  
... women are **less influential** ( $p < 0.001$ ).

[Link: Regressions](#)

## TREATMENT EFFECT ON WOMEN

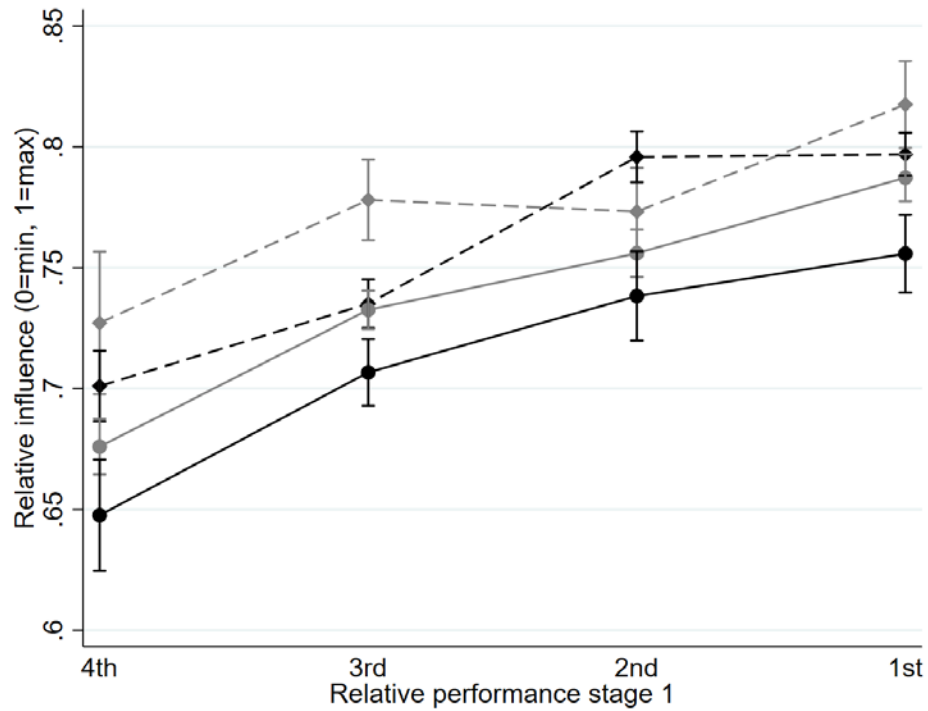
On average, controlling for relative performance, compared to women in female majority teams ...

... women in male majority teams are **less influential** ( $p = 0.015$ ).

[Link: Regressions](#)

- Men also speak more in the team discussions (28.2% vs. 22.7%).
- On average, men speak significantly more than 25% of the time ( $p < 0.001$ ), and women significantly less ( $p = 0.008$ ).
- However, we find no significant effect of team gender composition on women's average speaking time.

# INFLUENCE



---◆---  
Men in female majority teams

---◆---  
Men in male majority teams

—●—  
Women in female majority teams

—●—  
Women in male majority teams

# PERFORMANCE

---

**Do women in male majority teams perform worse than women in female majority teams?**

## **Individual performance in the second "Desert Survival" task**

- On average, controlling for relative performance in the first task...  
... there is no significant gender gap in performance in the second task (diff=1.005 points,  $p=0.090$ ).  
... there is no significant difference in performance in the second task between women in male majority teams and women in female majority teams (diff=0.333 points,  $p=0.736$ ).

# GENDER STEREOTYPES

---

**Do women in male majority teams face more biased team members than women in female majority teams?**

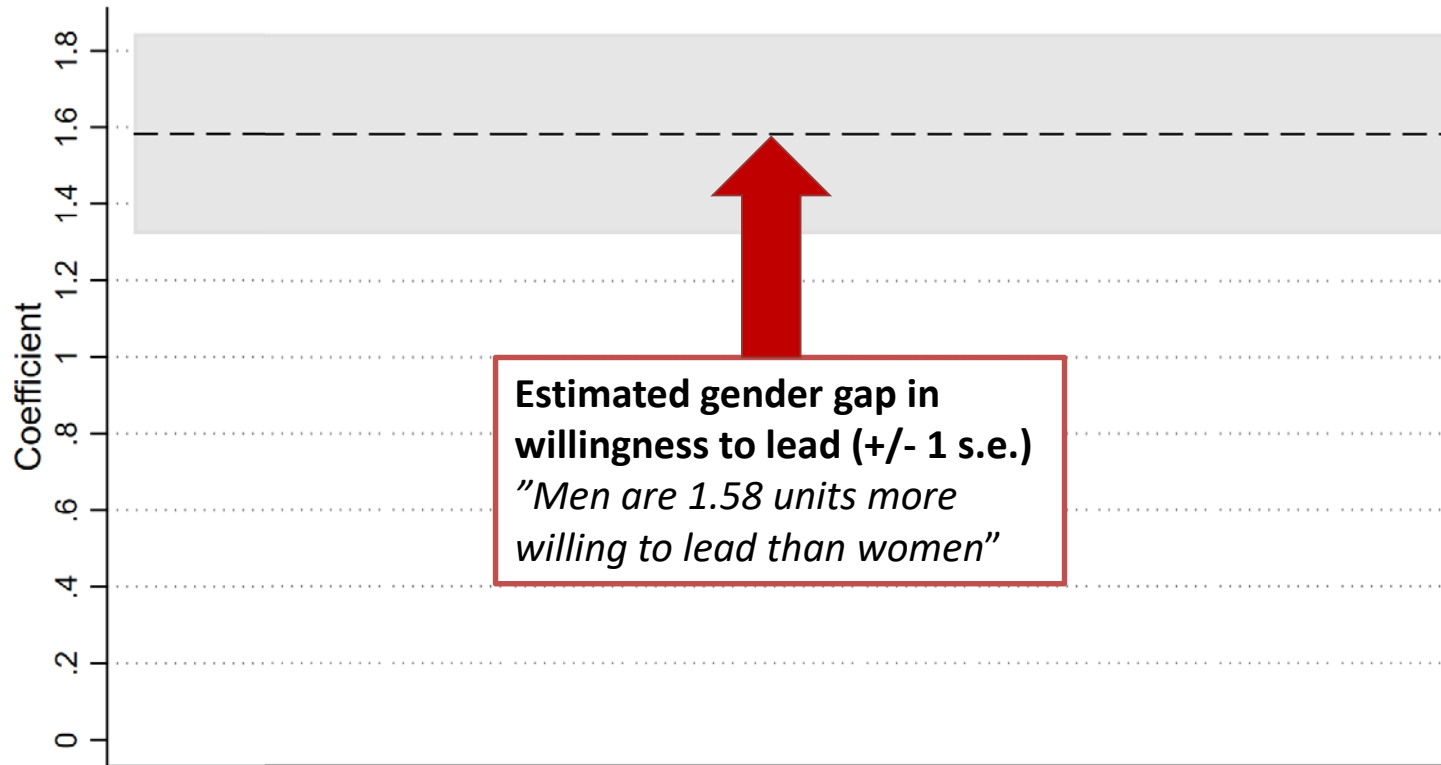
## **IAT score**

- After the experiment, all participants complete an Implicit Association Test (IAT), eliciting their implicit associations between being male and leadership.
- On average, controlling for relative performance in the first task...

**... women express less of a male-leadership association than men do** ( $p < 0.001$ ).

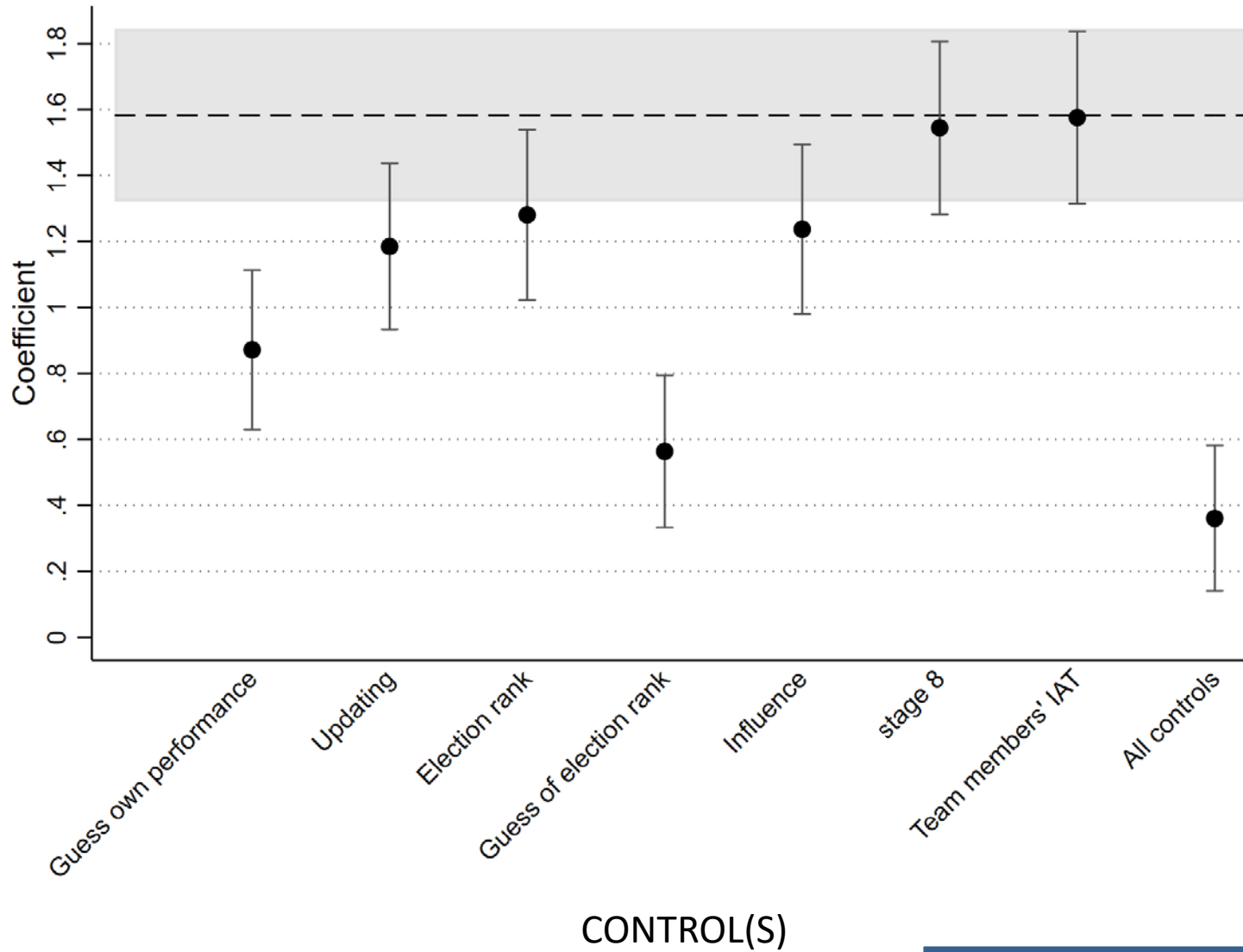
**... women in male majority teams face team members with a stronger male-leadership association than women in female majority teams** ( $p = 0.004$ ).

# ACCOUNTING FOR THE GENDER GAP

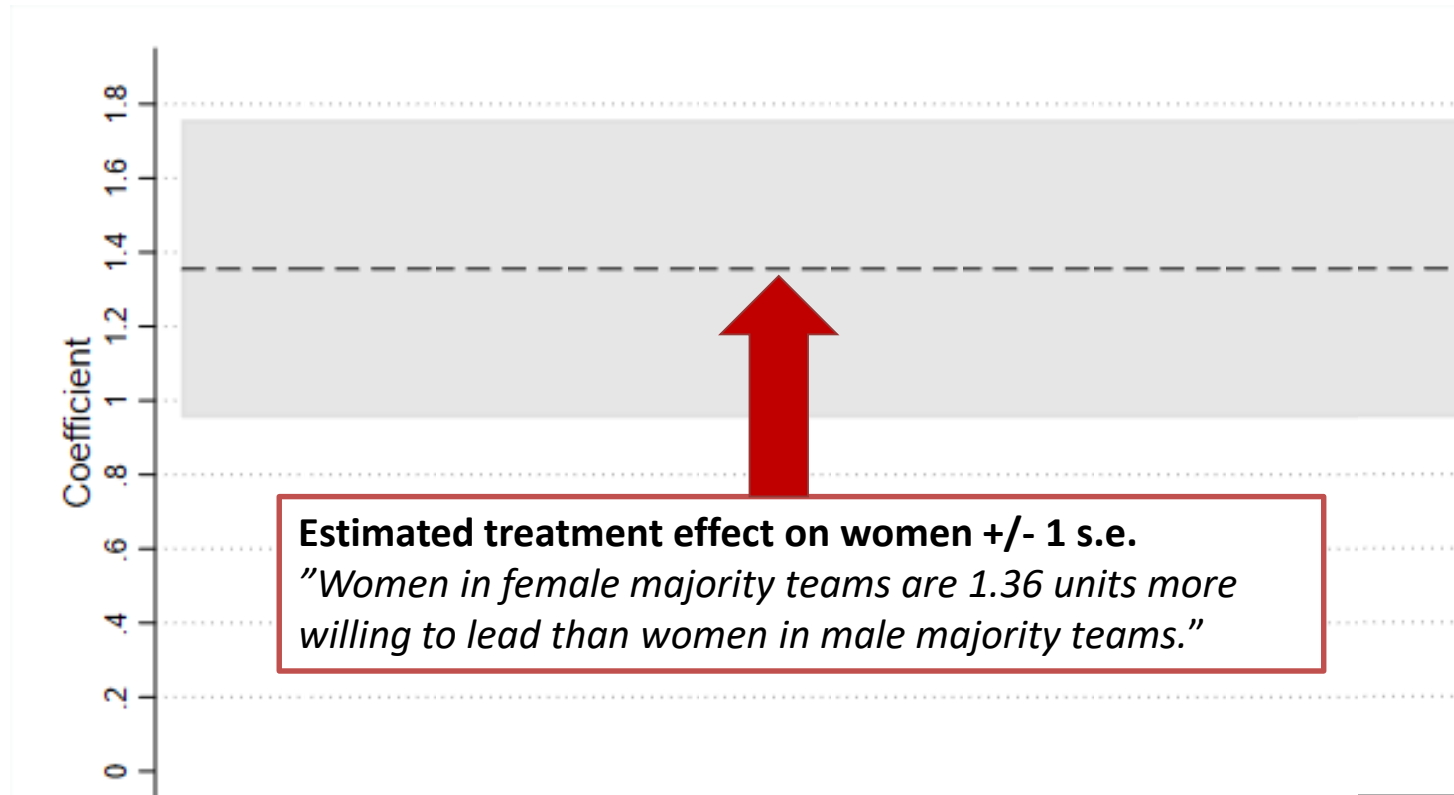


- How much does the estimated gender gap in willingness to lead decline when we control for these measures?

# ACCOUNTING FOR THE GENDER GAP

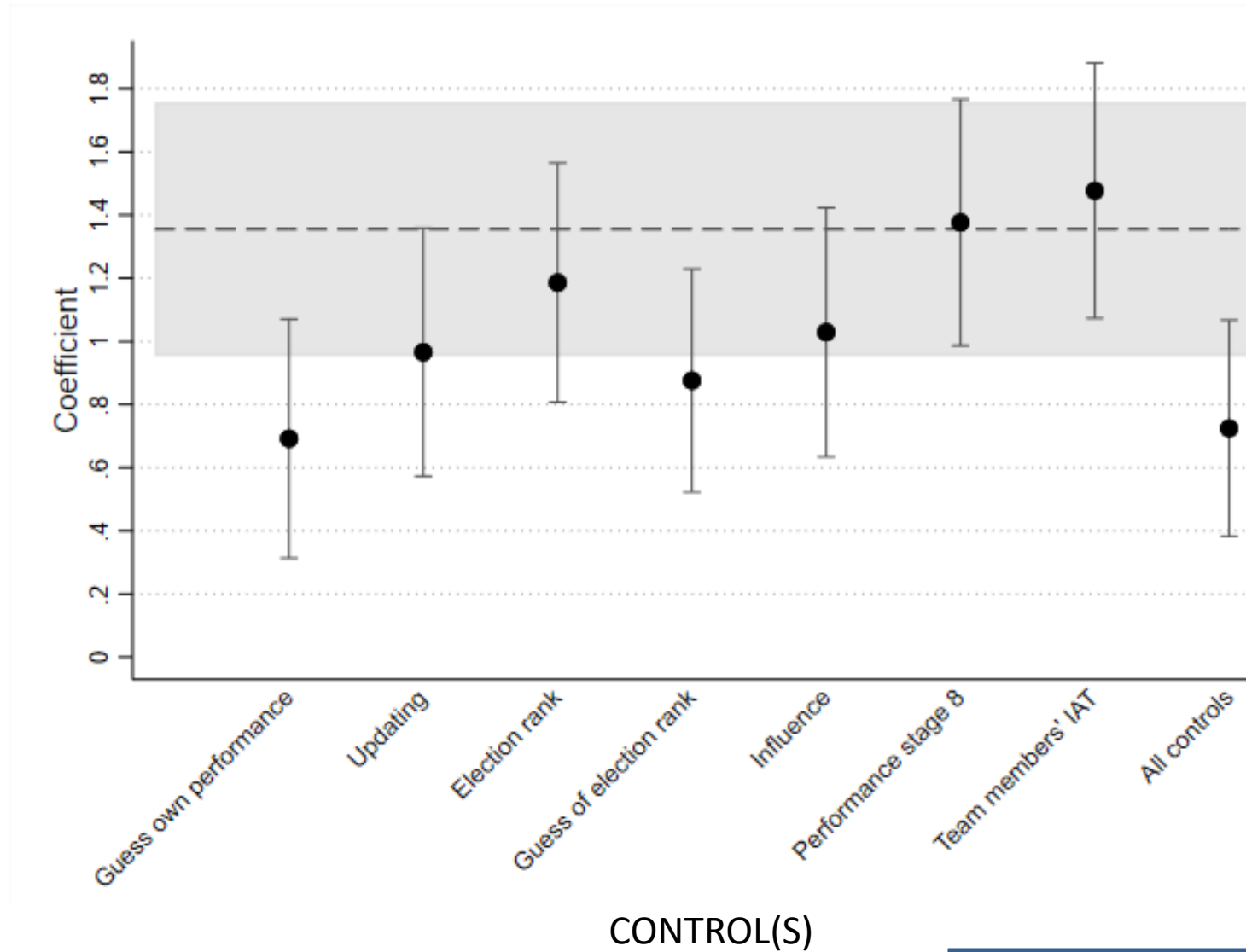


# ACCOUNTING FOR THE TREATMENT EFFECT



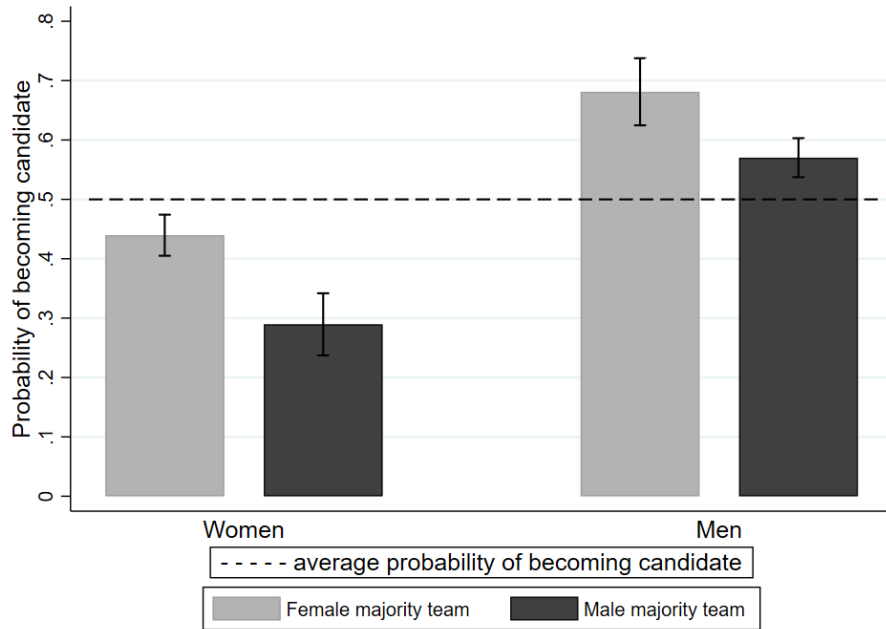
- How much does the estimated treatment effect on women decline when we control for these measures?

# ACCOUNTING FOR THE TREATMENT EFFECT

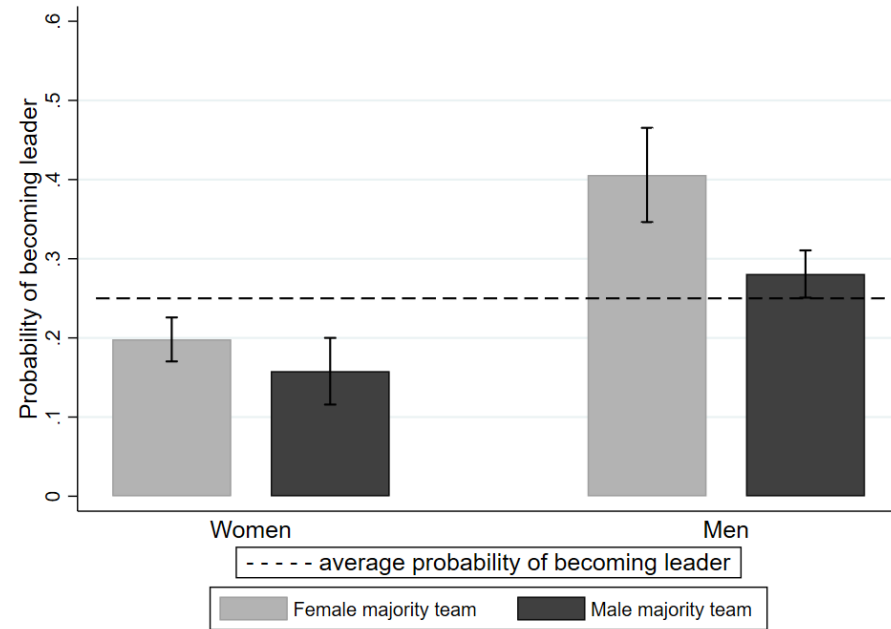


# PROBABILITY OF BECOMING CANDIDATE AND LEADER

## BECOMING CANDIDATE



## BECOMING LEADER



- Women are less likely than men to become candidates and leaders regardless of team gender composition.

# HOW WELL DO LEADERS PERFORM?

---

## MEASURE OF LEADERSHIP PERFORMANCE

- Leaders: Answer on behalf of the team for the "Desert survival" task (sees the suggested individual answers from all team members).
- Non-leaders: Identical task as the leader, performed at the same time, but unincentivized.

## DID THE TEAMS CHOOSE THE MOST ABLE LEADERS?

- In terms of this measure, **33% of teams chose the most able leader** (37% of male majority teams, 29% of female majority teams,  $p=0.315$ ).
- One reason behind the (non-significant) tendency for male majority teams to elect higher performing leaders appears to be that **female majority teams elect low-performing men too often.**

# SUMMARY AND CONCLUSIONS

---

- On average, men are more willing to become the group leader than women
  - Holds regardless of relative performance and gender group composition
- On average, women are **more willing to become the leader in female majority groups** than in male majority groups
  - Holds for all levels of relative performance
- **Relative performance beliefs and expected electoral support** seem to be particularly important in explaining these differences

# SUMMARY AND CONCLUSIONS

---

- Despite increased leadership motivation, women are **not more likely to become the leader of female majority groups**
  - Men are also more willing to become the leader in female majority groups (albeit not significantly so)
  - Men get a large share of the votes in female majority groups
- **Being in minority has a very different effect on men compared to women**
  - Men in female majority teams fare best in many regards, while women in male majority teams fare worst.
- Even though it does not translate into leadership, it is noteworthy that **women become significantly more confident and motivated to become leaders in female majority groups**
  - This is likely to correlate with other behaviors and feelings

# SUMMARY AND CONCLUSIONS

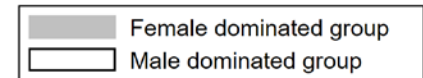
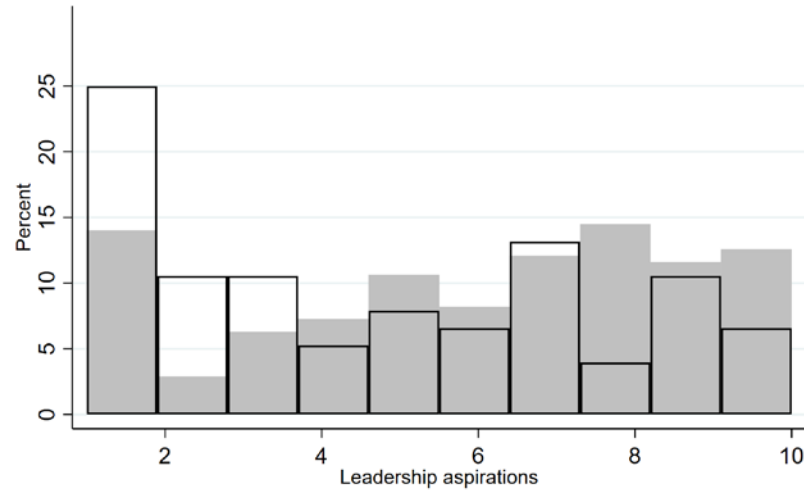
---

- The groups do not always pick the most able leaders
  - On average, the least able men are more willing to become the leader, and receive more votes, than the most able women
  - In such settings, maybe better to use other selection mechanisms
- More broadly our results, in addition to help understanding gender gaps in labor markets in several different ways, it also speaks to the current debate on gender quotas, how to structure team work, and how to attract, retain, and identify the most competent individuals to top positions and male dominated occupations

THANK YOU !

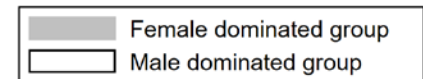
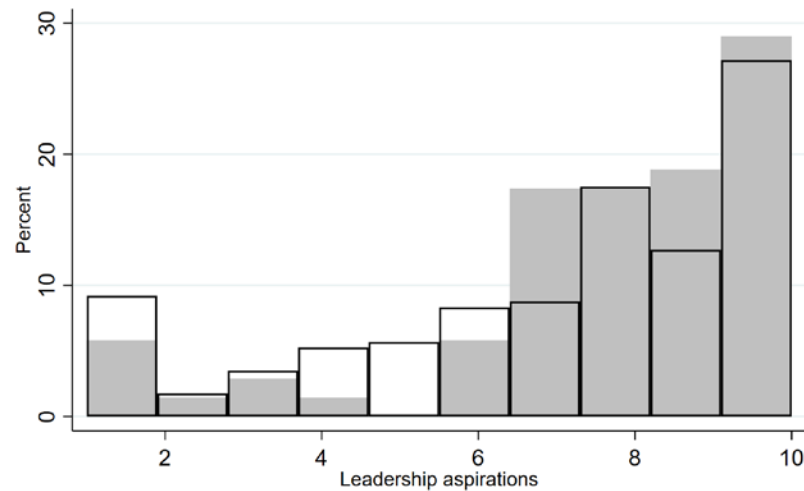
# DISTRIBUTIONS

## WOMEN: DISTRIBUTION OF LEADERSHIP ASPIRATIONS



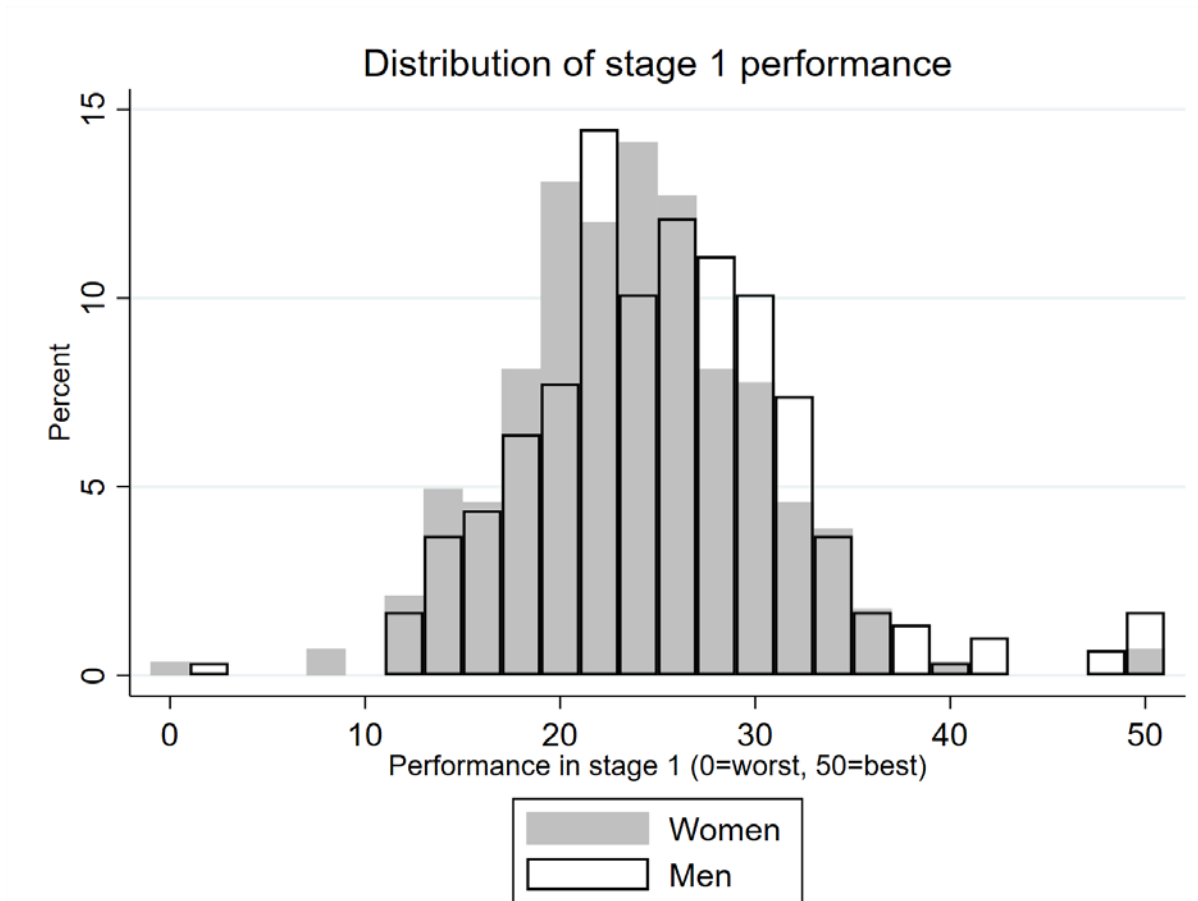
• KS-test:  $p < 0.01$

## MEN: DISTRIBUTION OF LEADERSHIP ASPIRATIONS



• KS-test: n.s.

# APPENDIX: GENDER DIFFERENCES IN PERFORMANCE



- On average, men score 1.81 points (0.26 SD) higher than women in the first task ( $p=0.0013$ )
- KS-test:  $p<0.05$

# APPENDIX: MAIN REGRESSION

**Table 2:** Differences in leadership aspirations across gender and team composition

	<i>Dependent variable: Leadership Aspiration (1-10)</i>			
	(1)	(2)	(3)	(4)
Female	-1.633*** (0.261)	-1.584*** (0.262)	-2.491*** (0.470)	-2.123** (0.657)
Female majority team			0.673 (0.355)	0.650 (0.353)
Female X Female majority team			0.713 (0.593)	0.706 (0.588)
Constant	7.266*** (0.166)	7.887*** (0.312)	7.110*** (0.195)	7.559*** (0.355)
N	580	580	580	580
<b>Controls:</b>				
Relative performance Stage 1		YES		YES
Female X Relative perf. Stage 1				YES
<b>F-test:</b>				
Female majority team + 'Female X Female majority team':			1.386*** (F=11.816)	1.355*** (F=11.377)

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

*Note:* OLS regressions using leadership aspirations (1-10) as the dependent variable. Standard errors are clustered at the team level. The final row shows results from an F-test, testing the treatment effect for women.

# APPENDIX: GENDER GAP IN MEDIATING MEASURES

**Table 3: Overall gender gaps**

Outcome variable:	Guess rank first task <i>1=worst</i> <i>4=best</i>	Updating <i>0=minimum</i> <i>1=maximum</i>	Rank election <i>1=lowest</i> <i>4=highest</i>	Guess rank election <i>1=lowest</i> <i>4=highest</i>	Influence <i>0=minimum</i> <i>1=maximum</i>	Performance second task <i>0=worst</i> <i>50=best</i>	IAT scores of team members <i>-2=female bias</i> <i>2=male bias</i>
Female	-0.441*** (0.066)	0.043*** (0.008)	-0.424*** (0.088)	-0.578*** (0.067)	-0.029*** (0.007)	-1.005 (0.589)	-0.014 (0.016)
Constant	3.304*** (0.087)	0.133*** (0.011)	3.027*** (0.114)	3.227*** (0.090)	0.846*** (0.009)	23.396*** (0.791)	0.216*** (0.019)
<i>N</i>	580	580	580	580	580	580	577

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

*Note:* Standard errors are clustered on the team level. All regressions are OLS regressions including controls for relative performance in Stage 1 (1=best, 4=worst). '*Guess rank Stage 1*' denotes the participant's guess (from Stage 4) of their within-team ranking in terms of performance in Stage 1 (1=worst, 4=best). '*Updating Stage 3*' indicates how much the participant updated their individual answer from Stage 1 in Stage 3, relative to the other team members (0=the participant did not update, 1=the participant was the only one in the team who updated). '*Rank election*' denotes the participant's rank in the election based on all votes (1=last, 4=first, equal observations are assigned the average rank). '*Guess rank election*' denotes the participant's guess of their rank in the election (1=last, 4=first). '*Influence*' denotes how close the team answer in Stage 2 was to the participant's individual answer from Stage 1, relative to the other team members (0-1, where 1=the team solution was identical to the participant's solution). '*Performance Stage 8*' indicates the participant's individual performance in the desert survival task in Stage 8 (0=worst possible, 50=best possible).

# APPENDIX: TREATMENT EFFECT ON MEDIATING MEASURES

**Table 4:** Effects of gender and team gender composition

Outcome variable:	Guess rank first task <i>1=worst</i> <i>4=best</i>	Updating Stage 3 <i>0=minimum</i> <i>1=maximum</i>	Rank election <i>1=lowest</i> <i>4=highest</i>	Guess rank election <i>1=lowest</i> <i>4=highest</i>	Influence <i>0=minimum</i> <i>1=maximum</i>	Performance second task <i>0=worst</i> <i>50=best</i>	IAT scores of team members <i>-2=female bias</i> <i>2=male bias</i>
Female	-0.658*** (0.180)	0.051* (0.022)	-0.292 (0.257)	-0.591** (0.216)	-0.034 (0.017)	-1.331 (1.624)	0.006 (0.036)
Female majority team	0.221 (0.113)	-0.027* (0.012)	0.346** (0.131)	0.463*** (0.113)	0.016 (0.010)	0.054 (1.046)	-0.074* (0.028)
Female X Female majority team	0.150 (0.161)	-0.005 (0.021)	-0.120 (0.231)	-0.201 (0.169)	0.007 (0.018)	-0.387 (1.398)	-0.014 (0.022)
Constant	3.239*** (0.121)	0.145*** (0.015)	2.837*** (0.150)	3.072*** (0.126)	0.838*** (0.012)	28.652*** (1.052)	0.248*** (0.024)
<i>N</i>	580	580	580	580	580	580	580.000
<b>F-test:</b>							
'Female maj. team' + 'Female X Female maj. team'	0.371*** (F=14.445)	-0.032** (F=7.937)	0.227 (F=3.100)	0.262* (F=5.549)	0.023* (F=6.100)	-0.333 (F=0.114)	-0.088** (F=8.587)

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

*Note:* Standard errors are clustered on the team level. All regressions are OLS regressions including controls for relative performance in Stage 1 (1=best, 4=worst), and relative performance in Stage 1 interacted with *Female*. The final row presents results from an F-test, testing the treatment effect for women. '*Guess rank Stage 1*' denotes the participant's guess (from Stage 4) of their within-team ranking in terms of performance in Stage 1 (1=worst, 4=best). '*Updating Stage 3*' indicates how much the participant updated their individual answer from Stage 1 in Stage 3, relative to the other team members (0=the participant did not update, 1=the participant was the only one in the team who updated). '*Rank election*' denotes the participant's rank in the election based on all votes (1=last, 4=first, equal observations are assigned the average rank). '*Guess rank election*' denotes the participant's guess of their rank in the election (1=last, 4=first). '*Influence*' denotes how close the team answer in Stage 2 was to the participant's individual answer from Stage 1, relative to the other team members (0-1, where 1=the team solution was identical to the participant's solution). '*Performance Stage 8*' indicates the participant's individual performance in the desert survival task in Stage 8 (0=worst possible, 50=best possible).

# APPENDIX: GENDER GAP CONTROLLING FOR MECHANISMS

**Table A2:** Changes in estimated gender difference after adding control variables

Control variable	Coefficient (overall gender gap)	Change compared to baseline
Baseline (no controls)	<b>-1.584***</b> (0.262)	
Guess rank Stage 1	-0.872*** (0.242)	-45%* (z=2.00)
Updating Stage 3	-1.185*** (0.252)	-25% (z=1.10)
Rank election	-1.281*** (0.259)	-19% (z=0.82)
Guess rank election	-0.564* (0.231)	-64%** (z=2.92)
Influence	-1.237*** (0.256)	-22% (z=0.95)
Performance Stage 8	-1.545*** (0.262)	-2.5% (z=0.11)
IAT score of other team members	-1.576*** (0.261)	+0.5% (z=0.02)
All control variables	-0.361 (0.220)	-77%*** (z=3.57)

*Notes.* Each reported coefficient is the coefficient of *Female* in a separate OLS regression using the following specification:  $Leadership\ aspiration = \beta_0 + \beta_1 Female + \beta_2 Relative\ Performance + \beta_3 Control\ variable$ . The bold coefficient in the first row is obtained using no control variables, and the coefficients reported in the subsequent rows are obtained controlling for the variable indicated in the first column. Standard errors are clustered on the team level. The reported change in the third column indicates the change in the coefficient of *Female* compared to the specification using no control variables.

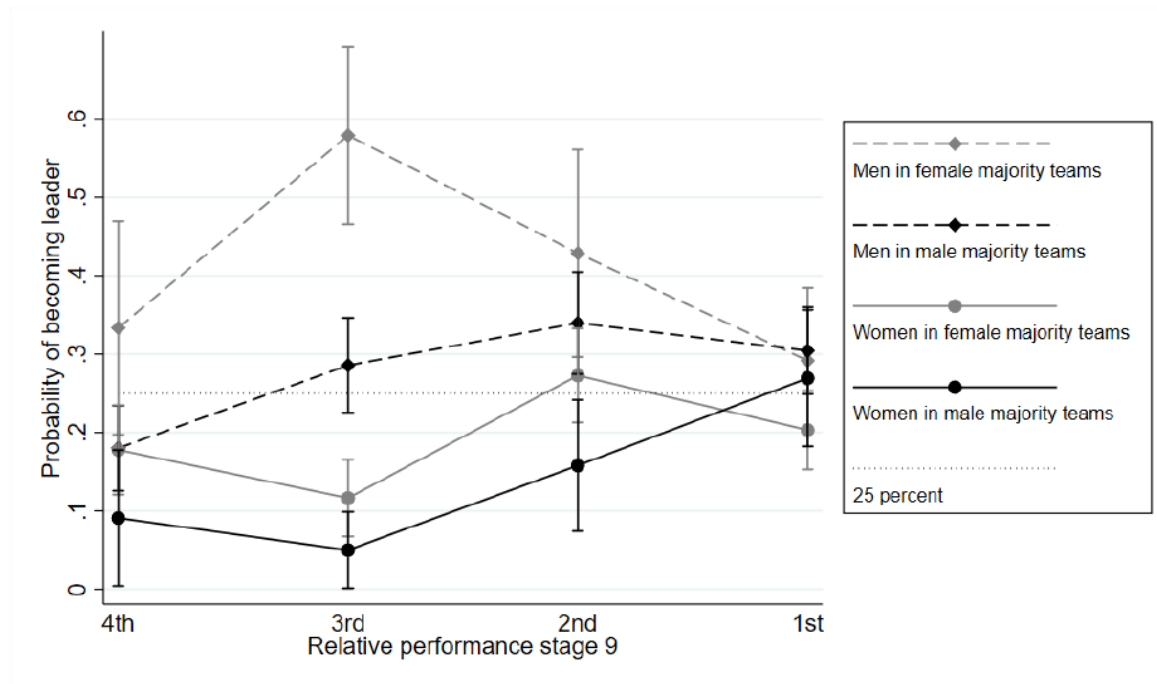
# APPENDIX: TREATMENT EFFECT CONTROLLING FOR MECHANISMS

**Table A3:** Changes in estimated treatment effect after adding control variables

Control variable	(a) WOMEN		(b) MEN	
	Coefficient (treatment effect)	Change compared to baseline	Coefficient (treatment effect)	Change compared to baseline
Baseline (no controls)	<b>1.355***</b> (0.402)		<b>0.650</b> (0.353)	
Guess rank Stage 1	0.692 (0.379)	-49% (z=1.20)	0.347 (0.341)	-47% (z=0.62)
Updating Stage 3	0.965 (0.392)	-29% (z=0.69)	0.480 (0.346)	-26% (z=0.34)
Rank election	1.186 (0.379)	-12% (z=0.31)	0.440 (0.350)	-32% (z=0.42)
Guess rank election	0.876 (0.353)	-35% (z=0.90)	-0.122 (0.338)	-119% (z=1.59)
Influence	1.029 (0.394)	-24% (z=0.58)	0.496 (0.352)	-24% (z=0.31)
Performance Stage 8	1.376 (0.390)	+1.5% (z=0.0)	0.649 (0.354)	-0.0% (z=0.00)
IAT score of other team members	1.477 (0.404)	+9.0% (z=0.21)	0.698* (0.338)	+7.4% (z=0.10)
All control variables	0.725 (0.342)	-46% (z=1.19)	-0.137 (0.319)	-121% (z=1.65)

*Notes.* All reported coefficients are obtained in separate OLS regressions using the following specification:  $Leadership\ aspiration = \beta_0 + \beta_1 Female + \beta_2 Team\ Female + \beta_3 (Female \times Team\ Female) + \beta_4 Relative\ Performance + \beta_5 (Female \times Relative\ performance) + \beta_6 Control\ variable + \beta_7 (Female \times Control\ variable)$ . Panel (a) reports the sum of the coefficients of *Team Female* and *Female X Team Female* (standard errors are obtained using STATA's 'lincom' command), while panel (b) reports the coefficient of *Team Female*. The bold coefficients in the first row are obtained using no control variables, and the coefficients reported in the subsequent rows are obtained controlling for the variable indicated in the first column. Standard errors are clustered on the team level. The changes reported in the third and fifth columns indicate the changes in the coefficients of interest compared to the specifications using no control variables.

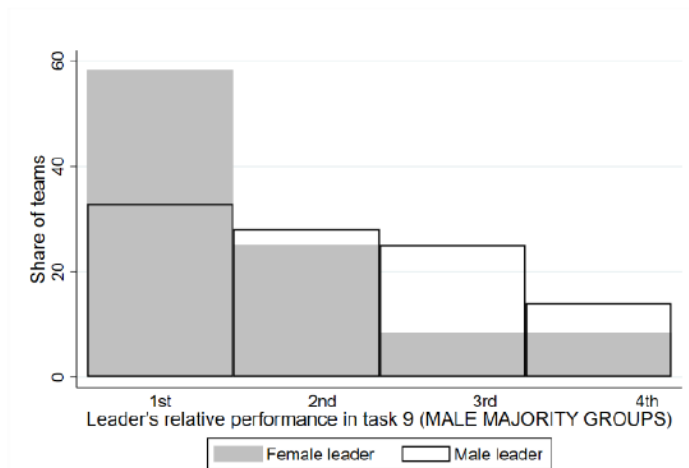
# APPENDIX: PROBABILITY OF BECOMING LEADER



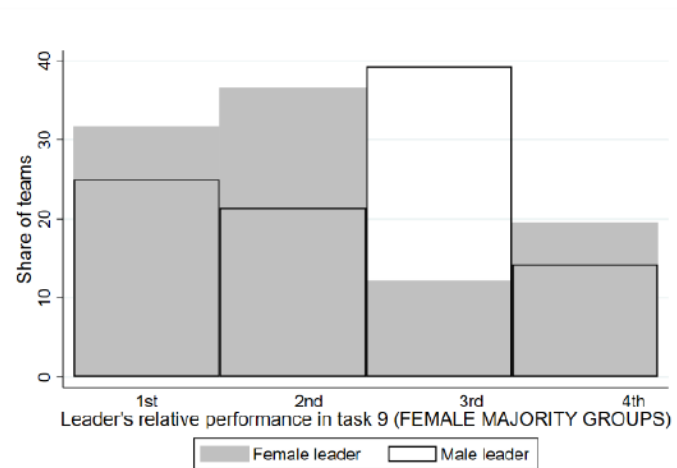
**Figure 9:** Probability of becoming the leader, by relative performance, gender, and treatment

*Note:* Error bars represent standard errors.

# APPENDIX: LEADER'S RELATIVE PERFORMANCE



(a) Male majority teams



(b) Female majority teams

**Figure A5:** Leader's relative performance in part 9

# APPENDIX: MEDIATING MEASURES

**Table A1:** Description of mediating measures

<b>Mediating Measure</b>	<b>Description</b>
<i>Guess rank (first task)</i>	The participant's incentivized guess of their performance ranking, elicited in step 4 of the experiment. Discrete variable ranging from 1 (worst) to 4 (best).
<i>Updating</i>	Sum (over all 10 items) of absolute difference in rank between the participant's original answer (from step 1) and their updated answer (from step 3), divided by the sum of all team members' differences. Continuous measure ranging from 0 (the participant did not update at all) to 1 (the participant was the only one in their team updating).
<i>Rank in election</i>	The participant's rank in the election. Discrete variable ranging from 1 (the participant was ranked last) and 4 (the participant was ranked first).
<i>Guess rank election</i>	The participant's incentivized guess of their rank in the election, elicited in step 7 of the experiment. Discrete variable ranging from 1 (the participant was ranked last) and 4 (the participant was ranked first).
<i>Influence</i>	Sum (over all 10 items) of absolute difference in rank between the participant's individual answer (from step 1) and their team's joint answer, divided by the sum of all team members' differences. Continuous measure ranging from 0 (the participant had no influence in the team's solution) to 1 (the team's solution was identical to the participant's individual solution).
<i>Performance (second task)</i>	Individual performance in the second task. Discrete variable ranging from 0 (worst possible solution) to 50 (best possible solution)
<i>Average IAT score (among team members)</i>	Average IAT score among team members, excluding the participant. Continuous variable ranging from -2 to 2, where a positive (negative) score indicates that the team members are biased in the sense that they find it easier (more difficult) to associate men with leadership than women.

# WHO BECOMES THE LEADER?

