It's Not About the Money – Or Is It? Stereotypes and the Gender Application Gap*

Paula Scholz**

February 28, 2025

Abstract

This study investigates how leadership responsibilities, salary differences, and gendered stereotypes influence the willingness to pursue leadership roles. Using a controlled laboratory experiment, we focus on communication and coordination responsibilities. In the experiment, subjects are randomly assigned to leadership positions in a public goods game in which the leader communicates with and coordinates the team. Afterwards, we elicit the willingness to pay for becoming a leader varying whether the position comes with a small or large salary increase. We find that women have a substantially lower willingness to pay to attain the leadership position if and only if it comes with a high salary. Despite women being equally effective team leaders as men, belief elicitation shows that high salaries shift leadership roles in our study from being perceived as stereotypical female to stereotypical male. This stereotypical perception of associating a highly paid leader with men translates into subjects' willingness to pay to attain the position. Exogenous exposure to leadership roles does not reduce the gender application gap, suggesting that experience alone cannot overcome instilled stereotypes.

*Acknowledgments: I especially thank my supervisor, Dirk Sliwka, for his valuable support throughout this project. Moreover, I thank Katherine B. Coffman, Ingrid Haegele, Saskia Opitz, Magdalena Smyk-Szymańska, Timo Vogelsang as well as participants of the Gender Gaps Conference 2024 in Warsaw, the PhD Course on Subjective Beliefs, Attention and Economic Behavior 2023 at the University of Copenhagen, the BMS Research Afternoon 2024, the YEP Junior Workshop in Gender and Labor Economcis 2024, the C-SEB Early Ideas Workshop, and the C-SEB Workshop 2024 in Cologne for helpful comments and discussions. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged. A pilot study was funded by the Center for Social and Economic Behavior in form of the C-SEB Student Award.

The experiment was approved by the Ethics Committee of the Faculty of Economics and Social Sciences (ERC-FMES) of the University of Cologne in July 2023 (Reference: 230038PT) and preregistered at the AEA RCT Registry with the ID *AEARCTR-0011984* on September 04, 2023. A pilot study was pre-registered under *AEARCTR-0010390*.

**University of Cologne, Department Corporate Development. Email: paula.scholz@wiso.uni-koeln.de.

1 Introduction

The representation of women in the labor market has increased substantially over the past decades, accompanied by a narrowing gender earnings gap (Goldin, 2014). However, gender disparities persist, especially in leadership roles. In 2024, only 34% of directors in Standard & Poor's (S&P) 500 firms were women (SpencerStuart, 2024).[1] While this inequality is most striking at the top, such gender disparities are also prevalent earlier in the career pipeline (McKinsey and LeanIn.Org, 2024; Azmat et al., 2024; Benson et al., 2024; Haegele, 2024; Bircan et al., 2024). Known as the 'broken rung', this phenomenon describes the sharp drop in the share of female employees advancing from entry-level roles to first-level managerial positions. With a reversed education gap and greatly reduced experience gap (Blau and Kahn, 2017), losing talented female workers early in their careers is inefficient as it reduces the size of talent pools available for promotions[2] and reinforces gender disparities in top positions.

A crucial driver of this broken rung is a gender gap in applications (Haegele, 2024; Fluchtmann et al., 2024). Different potential drivers of this application gap have been discussed in the literature. Salary expectations and career incentives influence application behavior differently by gender: women on average are more likely to take on low-promotability tasks which are not helpful for career advancement (Babcock et al., 2017), are less frequently assigned leadership roles ("high-promotability tasks") early in their careers (Bircan et al., 2024), and receive lower financial rewards for similar performance (Barker and Gil, 2023). Given that women tend to have lower self-confidence than men in professional settings (Exley and Kessler, 2022; Coffman et al., 2024), they may be particularly hesitant to apply for leadership roles without prior experience. Additionally, the persistent stereotype that successful leaders are male (Badura et al., 2018; Koenig et al., 2011) may further contribute to the application gap.

The key question we address in this paper is whether a mere shift in salary differences can affect the willingness to pay to attain a leadership position and trigger different gender stereotypes. Moreover, we test whether exogenous exposure to leadership experience can mitigate these effects.

We study these aspects in a lab experiment. The laboratory environment allows to i) isolate the effects of different salaries on application behavior while holding all other task dimensions fixed and ii) circumvent the self-selection problem when studying experience effects. In the experiment, we exogenously vary (i) whether a leadership role comes with a small or large salary increase and (ii) whether subjects have leadership experience at the application stage. The three-stage experiment is structured as follows: in Stage 1, participants are assigned to groups of three and play a standard public goods game. One participant in each group is randomly assigned as the team leader whose key task is to communicate with and coordinate the team. Before subjects decide about their contribution levels, the team leaders can communicate with their team members in bilateral chat

[1] Similarly, in Europe, the share of female board members in publicly listed firms stood at just 25% in 2017 (Smith, 2018).

[2] Maximizing the pool of qualified employees for leadership positions is crucial because already lower-level managers significantly impact subordinates' productivity (Lazear et al., 2015; Fenizia, 2022).

rooms. Team members can thus only chat with the team leader but not with the other team member, which makes the team leader the crucial player to communicate and coordinate the entire team. Leadership responsibilities in this stage are not remunerated. In the second stage, we elicit subjects' willingness to pay to attain a leadership role in Stage 3 using a second-price auction. Importantly, this leadership position comes with a large (50%) or small (10%) salary increase[3]. By exploiting the random assignment of leadership roles in Stage 1, we thus also can analyze the causal effect of leadership experience on the willingness to pay to attain a leadership position by gender.

We find that women have a lower willingness to pay to attain leadership positions when this position comes with a high salary. Interestingly, this gender gap is not observed for low-salary positions. Standard explanations such as differences in confidence, risk aversion, or leadership ability do not fully account for these findings. On the contrary, the data on team members' contributions rather suggest that women are at least as good as male leaders in our setting. In contrast to prior expectations, we do not see that the gender application gap for the high salary role is reduced by experience as the exogenous exposure to the leadership role in Stage 1 does not eliminate or reduce the gender gap in the willingness to pay for the leadership position. Unlike in Coffman et al. (2021, 2023), we thus do not detect any gender difference in the reactions to different experience or feedback when being a leader.

We furthermore find that the mere shift in salaries changes the perception of leadership roles from stereotypical female to male stereotyped. We also find a gender gap in a survey-based measure for the willingness to claim the lead (Giessner et al., 2022), i.e. women express a lower general willingness to actively pursue leadership roles. Importantly, this willingness to claim the lead is more strongly associated with the willingness to pay for high salary leadership positions than for low salary leadership positions.

This study contributes to several strands of the literature. First, our study adds to the literature on leadership stereotypes (e.g. Badura et al., 2018; Koenig et al., 2011) by showing that the stereotype of male leaders is induced by a mere rise in salary in our setting. At the same time, we find a gender gap in application aspirations for high salary positions only. Coffman et al. (2024) showed that women are less likely to apply for higher return but more challenging work in a male-typed domain. In our experiment, women exhibit a lower willingness to pay compared to men for assuming a leadership position *although* the leadership responsibility, task complexity, and other task dimensions remain the same. The only difference is the increase in salary, making it even more important to counteract the instilled stereotypes.

Second, we contribute to the literature on leadership training (Day, 2000; Martin et al., 2021) and the role of experience (e.g. Haegele, 2024; Bircan et al., 2024) by showing that exogenous exposure to leadership roles may have only limited effects on future application behavior. In our setting, first experience with being a leader did not reduce the gender gap in the willingness to invest to attain leadership positions.

[3]The low salary adds 10% while the high salary adds 50% to the fixed endowment as a risk-free earning in Stage 3. By using the strategy method, subjects indicate their willingness to pay for both positions.

Third, our study departs from the traditional 'leading by example' framework often employed in experiments, where leaders act primarily as first movers in public goods settings (e.g. Gächter et al., 2012; Güth et al., 2007; Eisenkopf, 2020; Gächter and Renner, 2018; Cappelen et al., 2016). Instead, we focus on the communication aspect of leadership[4]. Effective leaders spend much of their time building trust and proactively engaging with team members (McKinsey and LeanIn.Org, 2023), with CEOs dedicating approximately 70% of their time to interpersonal interactions (Bandiera et al., 2020).

Our study further relates and adds to the extensive literature on other drivers of the gender representation gaps, including for example gender biases in the hiring process (Carlsson and Eriksson, 2019; Bohnet et al., 2016), inflexible working times or missing part-time opportunities (Goldin, 2014), lower self-promotion of skills by women (Exley and Kessler, 2022), the lower willingness of women to enter competitive environments and their lower self-confidence compared to men (Niederle and Vesterlund, 2007) as well as lower social confidence, i.e. the willingness to publicly perform a task (Alan et al., 2020). The controlled environment provides the opportunity to rule out that the effects of salary-driven changes in stereotypes on the willingness to become a leader are induced by these other common aspects presented in the literature.

The remaining of this paper proceeds as follows. Section 2 introduces the experimental design. After presenting the data in Section 3, the main results are presented in Section 4. Mechanisms are discussed in Section 5. Section 6 concludes.

2 Experimental Design

The structure of the experiment can be divided into three main stages. In the first stage, subjects are stratified by gender and assigned to the leader position or team member position. Each group hence consists of one team leader and two team members. In these groups of three, subjects play a standard public good game (PGG). The marginal per capita return is 0.5. Subjects are endowed with 3.50€. This yields the following payoff function:

$$x_i = 3.5 - c_i + 0.5 \cdot \sum_{k=1}^{3} c_k \tag{1}$$

with c_i being player i 's contribution to the public good. Thus, subjects are confronted with a social dilemma situation: it is individually optimal to contribute nothing, i.e. to free-ride, because this will maximize the individual's payoff whereas it is socially optimal to contribute the entire endowment as this behavior maximizes the team's overall payoff. The PGG is adapted to the context of interest by adding a communication tool, i.e., a free-form chat that one player in the role of a leader can use to persuade others to contribute to the public good. The leaders can chat with *both* team members *simultaneously* in two *separate* chat rooms for three minutes before all subjects

[4]Eisenkopf (2014) investigates the impact of managers with communication power on competitive behavior of a group. Weber et al. (2001) and Antonakis et al. (2022) study the effect of leaders' speeches on a group's behavior.

decide about their contribution levels. The team members can only chat with the team leader but not with the other team member of their group. This approach is similar to the experiment of Eisenkopf (2014) in which managers are also characterized by communication power. The leader is thus the crucial player to achieve coordination and communication in the entire team. Furthermore, during the chat, team leaders have the possibility to persuade their team members to contribute their endowment. In this context, the leaders have an incentive to persuade the team members to contribute much to the public good - independent of their own action. Subjects are not allowed to reveal their identity or gender when chatting. The chats close after three minutes and subjects make their contribution decision simultaneously.

After the contribution decisions, subjects receive noisy information on the groups' overall contribution level adjusted by a randomly chosen integer between -2 and 2. Adding this noise term n, exogenously varies whether conditional on performance, i.e. the contribution level of the team members, leaders have made a positive ($n = \{0, 1, 2\}$) or negative ($n = \{-1, -2\}$) experience. These signals are capped to stay within an interval between 0€ and 10.50€ [5].

In the second stage, we measure participants' willingness to pay for becoming a leader in Stage 3 of the experiment, which again follows the same PGG protocoll as in Stage 1. Analogous to standard job search models (Cahuc et al., 2014) in which more effort increases the probability to get the job as well as to measure preferences to become a leader in an incentive-compatible way, applicants apply in form of a second price sealed bid auction (Vickrey, 1961). To apply to the leader position in Stage 3, subjects are endowed with an extra budget of 3.50€ independent of the payoffs they have earned in Stage 1. Different to Stage 1, leaders in Stage 3 receive a salary. This salary is either low (0.35€) or high (1.75€)[6]. We use the strategy method and elicit subjects' bids for both positions and it is randomly determined which bid is relevant for the remainder of the experiment. Each subject competes against two other subjects. Groups are randomly re-mixed after the first stage but will stay the same in the subsequent Stage 3.

Stage 3 is almost equivalent to Stage 1 as it again consists of a PGG with chats and noisy feedback. However, team leaders are not randomly assigned to their roles but leader positions are filled according to the application outcome[7].

In each stage, subjects receive the same endowment. It is randomly determined whether subjects' payment is based on Stages 1 and 2 or on Stages 2 and 3. At the beginning of the experiment, subjects' risk preferences are elicited using urn gambles. Furthermore, subjects are asked to state their relative persuasion ability compared to others, their confidence in leading a group, their reasons for (not) applying (open text) as well as their willingness to claim or grant the lead (Giessner et al., 2022). In the end of the experiment, stereotypes connected to this role were elicited both in an incentivized and non-incentivized way.

[5]For 238 subjects (26%), the noise term had to be capped. 235 noise terms were capped at the maximum, 3 at the minimum contribution level.

[6]Subjects are endowed with 3.50€ in the third stage. The low salary adds 10% while the high salary adds 50% to the fixed endowment as a risk-free potential earning in Stage 3 .

[7]That is, in each group, the participant with the highest bid will be the leader of the group in Stage 3 and pays the second-highest bid submitted in the group.

Throughout the experiment, subjects do not know the gender of the other participants. Since the experiment was conducted online, participants were also unaware of the gender distribution in their session

3 Data

The software oTree (Chen et al., 2016) was used for programming the experiment. The experiment was run online via the Cologne Laboratory of Economic Research from September 2023 to February 2024. In consequence, 71% of the subjects are students. Subjects were invited using the recruitment system ORSEE (Greiner, 2015) and participated online (not in person in the laboratory). The number of participants in the final sample is 901 of which 301 subjects are assigned to the team leader position in Stage 1 and 600 subjects are assigned to the team member position in Stage 1.

The average participant is 28 years old. 51% of the subjects are female and 71% of the sample are students with most of them (about one third) studying in the field of management, economics and social sciences. Subjects were stratified by gender and randomly assigned to the leadership position. Table A1 in the Appendix shows that the treatment and control groups are balanced with respect to gender, age, and whether subjects are students or employees. There are some imbalances with respect to the field of study as well as the type of occupation. All regressions will thus control for those variables.[8]

4 Gender Application Gap

As a first step, we measure men's and women's willingness to become a leader in Stage 2 of the experiment – the application phase. As laid out in the above we elicited their willingness to pay for a low and a high salary position using the strategy method.

Figure 1 depicts the mean bids of male and female subjects for the low (lhs) and high (rhs) salary position.[9] For the low salary position, we find (if anything) a small reversed gender application gap with women submitting 5% higher bids than men (0.41€ instead of 0.39€). For the high salary position, this pattern changes: men make 12.6% higher bids than women (1.25€ instead of 1.11€).[10]. Both men and women are more willing to become a leader if it is linked to a high salary

[8]Further, the gender composition of the sessions varied so that the gender compositions of the groups in Stage 1 of the experiment differ. We will therefore control for the gender composition of the groups in Stage 1. Subjects were not aware of the gender composition of their session or group.

[9]Figure A1 in the Appendix shows the distribution of bids for the low and high salary position separately for men and women.

[10]The average bid for the low salary position (0.40€) exceeds the offered salary by 5 cents. The average bid (1.18€) for the high salary position does not exceed the wage of 1.75€. Bids of 0.34€ and 1.74€ would ensure a positive gain of 1 cent. This assumes that subjects do not have other costs of assuming the positions. As the mean bid for the high salary position is much lower than this bid, it can be presumed that subjects perceive unobservable costs of being a leader with a high salary. This can be interpreted as a first suggestion that they only perceive the high salary position as a leadership position in which (costly) engagement (e.g. in form of taking on the responsibility of convincing the team members) is expected or required. The low salary position does not seem to be perceived as such as indicated by the average bid.

Figure 1: Mean Bids of Men and Women

but men are more eager to take over this position than women.

To investigate this further, we run OLS regressions of the following form:

$$y_{iS} = \alpha + \beta\, Female_i + \gamma\, S + \delta\, Female_i \times S + \eta Leader_i + \theta X_i + \varepsilon \tag{2}$$

with $Female_i$ being a dummy indicating the gender of person i, S an indicator taking a value of 1 for the high salary position, and $Leader_i$ an indicator taking a value of 1 if the subject was assigned to the leader position in Stage 1. X_i is a vector of control variables.[11] The dependent variable y_{iS} represents the submitted bids (ranging from 0 to the maximal bid of 3.5) of person i for salary level S. In each specification, standard errors are clustered at the Stage-1-group level to account for the correlation of the submitted bids within a cluster, i.e. a group in Stage 1[12].

Column (1) in Table 1[13] shows the results for the main specification as defined in Equation (2).

[11] For the main regressions, (seven) field of study dummies as well as (eight) occupation dummies will be included to account for the imbalances in these variables between the team leaders and team members. Furthermore, a categorical variable accounts for the gender composition of the group in Stage 1.

[12] Abadie et al. (2023) suggest that clustering at the individual level would suffice for the pooled specification while no clustering would be necessary for specifications (2) and (3) in Table 1 because the randomization unit is the individual. Following their suggestion (see Table A2 in the Appendix) does not change the significance of the results.

[13] Table A3 reports the specifications without controls and with controls including age as an additional control variable. The results are similar. 55.19% (51.45%) of female leaders (members) and 70.75% (63.57%) of male leaders (members) submitted a bid larger than zero for the low salary position. 75.97% (71.84%) of female leaders (members) and 87.07% (80.41%) of male leaders submitted a bid larger than zero for the high salary position. See Table A4 for results regarding this extensive margin.

Table 1: Gender Application Gap

	(1) Bid	(2) Bid High Salary	(3) Bid Low Salary
Female	-0.00293	-0.185**	-0.00373
	(0.0533)	(0.0857)	(0.0610)
High Salary	0.863***		
	(0.0358)		
Female × High Salary	-0.162***		
	(0.0516)		
Leader Experience	0.0333	0.0472	-0.0119
	(0.0515)	(0.0873)	(0.0570)
Female × Leader Experience		0.0172	0.0443
		(0.135)	(0.0854)
Constant	0.595**	1.468***	0.578*
	(0.283)	(0.285)	(0.307)
R-squared	0.204	0.0275	0.0151
No. obs.	1802	901	901

Note: The table reports OLS regressions with the height of the submitted bids as outcome variables. Column (1) reports the pooled results, Columns (2) and (3) split the sample into bids for the high and low salary positions, respectively. Clustered standard errors at the Stage 1 group level in parentheses. All specifications control for field of study, occupation, the gender composition of the groups. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Columns (2) and (3) refer only to bids for the high and low salary position, respectively. Column (1) shows that there is no significant gender gap in bids for the low salary position. For the high salary position, both men and women submit higher bids as compared to the low salary position but men's bids increase significantly and economically more than the ones submitted by women. Thus, there is a significant gender application gap for high salary positions (Column 2). This finding is in line with the observed gender gaps for top positions (e.g. McKinsey and LeanIn.Org, 2024), and women's higher likelihood of taking on low promotable tasks which typically come with lower pay (Babcock et al., 2017; Bircan et al., 2024). In our study, we can show that the mere increase in salary causes the emergence of a gender application gap for the highly paid position.

All specifications control for being assigned to the leadership position in Stage 1. There is no evidence for the leadership experience to affect the bidding behavior and thus the willingness to become a leader. As can be seen in Columns (2) and (3), we also do not find evidence that experience matters for women.

To sum our main results up, men are more eager than women to take over leader positions with high salary while we do not find evidence for such a gender gap if the leader's salary is low. Hence, the mere change of the salary level associated with a position holding all other factors constant

leads to the emergence of a gender application gap. This finding allows to conclude that it is not the leadership role per se that causes the application gap. If this were true, there should be a significant gender gap for both positions. However, the descriptive data even suggests that women are – if anything – more willing to become a leader with low salary compared to men. In the following, we will investigate the underlying mechanisms which explain why the large salary increase causes a gender application gap for leadership positions.

5 Mechanisms

The two leadership positions only differ in the level of the salary for the leader.[14] To understand the underlying reason for the change in the willingness to invest for leadership position which is induced by the increase in salary, this section explores which other factors change through the high salary. This analysis will also allow to learn more about implications for policies targeting the reduction of the observed gender application gap for high salary positions.

5.1 Leader Performance

To start with, it is important to understand whether the gender gap in the willingness to invest for leadership positions reflects true performance differences between male and female leaders. As leaders earn more when their teams are more successful, such a performance difference could potentially rationalize a gender gap in applications for the high salary position.

Table 2: Leaders' Contributions

	Stage 1	High Salary	Low Salary
	(1)	(2)	(3)
Female	0.217	0.318	0.339*
	(0.134)	(0.204)	(0.196)
Constant	2.958***	3.986***	3.098**
	(0.224)	(0.796)	(1.248)
R-squared	0.0415	0.172	0.107
No. obs.	301	124	165

Note: The table reports OLS regressions with the team leaders' contributions as outcome variables. Column (1) refers to Stage 1 leaders, Columns (2) and (3) to leaders in the high salary position and the low salary position, respectively. All regressions control for the gender of team members, field of study and occupation. Robust standard errors in parentheses. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

A key factor in assessing leadership performance is the leader's own behavior, measured by their individual contributions. As shown in Table 2, if anything, female leaders tend to contribute more to the public good game than their male counterparts. While the difference is statistically

[14] And indeed, this salary is mentioned most often when subjects explain their decision to apply in open-text fields.

significant only for leaders earning a low salary, the overall pattern of coefficients suggests a similar tendency across salary levels.

Beyond individual contributions, a leader's performance must also be evaluated through their team's performance, i.e. team members' contribution levels. It is the goal of the leader to persuade their team members to contribute their entire endowment. Regardless of the leaders' own actions or their salary, higher contributions of the team members will also increase the leaders' own payoffs. Therefore, even when receiving a high salary, leaders should aim for high team members' contributions. Leaders in Stage 1 have been *randomly* assigned to the position of a leader and thus the performance of male and female leaders in Stage 1 is informative about gender differences in performance as leaders in this specific task.

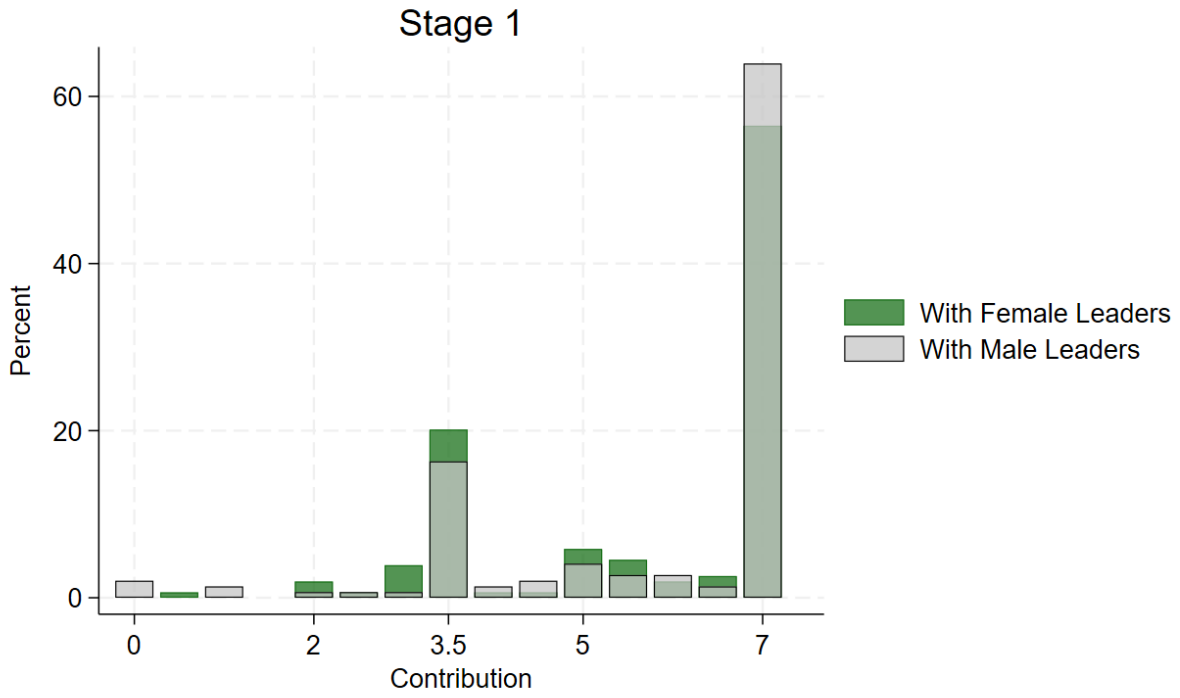

Figure 2: Contributions of team members in Stage 1

Figure 2 depicts the distribution of the team members' contributions in a group led by a man or a woman in Stage 1.[15] A Kolmogorov–Smirnov test could not reject equality of distributions between male and female leaders' performances (p=0.798). This indicates that there is no evidence for men and women to differ in their performance if in a role of a leader.

To further investigate gender differences in performance, we run OLS regressions considering additional factors. First, leaders' own contributions may affect their success as a leader as leaders seem to use statements about their own contributions as a tool to convince their team members[16]. Second, the performance of the leader is not purely driven by the leaders themselves but also

[15] See Figure A2 in the Appendix for an overview about distributions in all three Stages.

[16] According to the chat protocols, all team leaders in Stage 1 and Stage 3 discuss their contribution levels.

depends on the group's team members. Leaders did not know the gender composition of their group but there might be differences in men's and women's behaviors as team members. Therefore, we further control for the gender of leader's team members. Leaders can either communicate with two women, two men or a group of mixed gender[17].

Table 3: Leader's Performance

	Stage 1		Low Salary		High Salary	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.133	-0.197	-0.0797	-0.213	-0.0453	-0.189
	(0.208)	(0.207)	(0.357)	(0.358)	(0.339)	(0.325)
Own Contr.		0.294***		0.356**		0.452***
		(0.112)		(0.165)		(0.168)
Constant	7.521***	6.652***	8.020***	6.931***	3.985***	2.185**
	(0.369)	(0.523)	(0.659)	(0.834)	(0.528)	(0.863)
R-squared	0.0337	0.0661	0.0991	0.135	0.0937	0.166
No. obs.	301	301	164	164	124	124

Note: The table reports OLS regressions with the sum of team members' contributions in a group as outcome variables. Columns (1) and (2) refer to Stage 1 leaders, Columns (3) and (4) to leaders in the low salary position and Columns (5) and (6) to leaders in the high salary position. All regressions control for the team members' gender, the leaders' field of study and occupations. Robust standard errors in parentheses. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table 3 shows the results of these regressions separately for leaders in Stage 1 (No Salary), and Stage 3 (Low or High Salary). Columns (1), (3), and (5) confirm the missing significance in a gender difference among men and women as leaders. Adding the own contribution level shows that the leader's contribution level significantly and positively relates to the team members' contributions. Though this variable is endogenous, the result confirms many other studies on this relationship (Fischbacher et al., 2001; Fehr and Gächter, 2000; Gächter et al., 2012). It further shows that women are at least as good leaders in Stage 2 as compared to men. Even a negative point estimate does not speak against this statement: female leaders contribute (if anything) more than male leaders (see Table 2) and their own contributions are significantly related to higher contributions of the team members[18]. In all specifications, the female indicator remains insignificant. It can thus be concluded that neither those who were randomly selected for leadership nor those who self-selected into the positions show a significant gender gap in performance. Hence, there is no evidence for the gender application gap to reflect true performance differences.

[17]Mixed gender groups also include teams which have at least one team member who indicated "diverse" as gender (less than 1% of the sample).

[18]The larger contributions by female leaders do not significantly relate to the groups' overall contribution level: Table A11, Columns (1) – (3), in the Appendix can show that the overall group contribution is not significantly different for male and female leaders. Also, male and female leaders do not differ in their earnings received through the public good game (see Table A11, Columns (4) – (6), in the Appendix).

5.2 Risk Attitudes and Confidence

In a second-price auction, which is used as the application format in this study, risk preferences should not impact the bidders' actions in equilibrium because the weakly dominant strategy is to bid the true valuation independent of the bidders' risk preferences (Chen et al., 2013). While the theory states that there is no reason to believe that risk preferences matter, the scarce empirical evidence on gender differences in bidding in second-price sealed-bid auctions shows mixed results.[19] Therefore, we investigate the role of risk attitudes in our setting in which men and women compete against each other in form of an auction.

Many studies stress that men are more likely to enter competition because of different preferences beyond overconfidence and risk attitudes (e.g. Niederle and Vesterlund, 2007). However, more recent literature challenges these findings and points out that – if correctly measured – the entire gender gap in entry for competition can be explained by those personality traits (see e.g. Gillen et al. (2019)).

Prior confidence was elicited before subjects have been assigned to their roles and is thus not affected by the treatment. In particular, subjects answered four questions which in combination elicit how much subjects think they would be able to make a team member contribute more compared to a situation in which those were led by another team leader (ranging from -3.5 to 3.5)[20]. Positive (negative) values represent individuals who believe they are able to convince others to contribute more (less) than with another leader. This self-confidence measure is standardized to have a mean of zero and a standard deviation of one.[21] In addition, risk attitudes were elicited twice in the beginning of the study using two similar urn gambles[22]. These risk measures are also standardized to ease comparability.

Firstly, as found in many other studies, the average female participant in our study is on average more risk averse than the male but the mean difference is not significant at any conventional level (see Table 4). Secondly, the average man is also more confident than the average woman (not significant) in their ability to persuade their team members compared to another participant in a leader position.

[19] Women seem to bid less than men in a Vickrey auction for wine (Lecocq et al., 2005) and for a high-quality, limited-edition poster (Onderstal, 2020) but do not differ significantly in their bids for gourmet-chocolate (Rutström, 1998). While those analyses on gender differences may be affected by a gender difference in the preference for the goods auctioned, Chen et al. (2013) investigate the bidding behavior in a second-price sealed bid auction for an abstract object. Across the value distribution (see p. 38 of the online appendix on Yan Chen's webpage) and on average, Chen et al. (2013) do not find significant gender differences in bids for the abstract object for which they exogenously vary the value participants assign to this object.

[20] Subjects were asked to guess the average contribution of a groups' team member if 1) they themselves were the team leader and 1a) were able to chat or 1b) were not able to chat and 2) someone else was the team leader who 2a) was able to chat or 2b) was not able chat with the team members.

[21] We did not elicit prior confidence twice before treatment assignment because there was no reasonable way of implementing a second measure without anticipating anchoring bias which is advised to avoid by Gillen et al. (2019).

[22] We elicit certainty equivalents with two multiple price lists. Subjects could choose between a safe option and an urn game with 50% probability of winning. In one urn gamble, subjects could choose between a safe option ranging from 0€ to 1€ in increments of 10 cents and an urn gamble with 50% probability of winning 1€. In the other, the list covered safe options up to 1.50€ and subjects could win 1.50€ in the urn game. The order of lotteries was randomized.

Table 4: Gender Differences in Risk and Confidence

	(1) Risk1 (Std.)	(2) Risk2 (Std.)	(3) Prior Confidence (Std.)
Female	-0.0256	-0.0828	-0.0953
	(0.0670)	(0.0655)	(0.0713)
Constant	-0.142	0.542***	-0.0522
	(0.365)	(0.121)	(0.133)
No. obs.	901	901	901

Note: The table reports OLS regressions with two standardized risk measures in Columns (1) and (2) and the standardized measure of prior confidence in Column (3). Robust standard errors in parentheses. All specifications control for field of study and occupation.* for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

To reduce the bias introduced by measurement error in the risk measures, we employ the ORIV instrumentation technique developed by Gillen et al. (2019). As we have two risk measures, we can instrument the first risk measure with the second and vice versa. Taking the average estimate from the two instrumentation approaches reduces the attenuation bias caused by measurement error (Gillen et al., 2019).[23]

Table 5: Risk and Confidence

	(1) Bid High S.	(2) Bid Low S.
Female	-0.175**	0.0152
	(0.0718)	(0.0520)
Leader Experience	0.0569	0.0116
	(0.0677)	(0.0449)
Risk (Instr.)	0.0520	0.0686*
	(0.0570)	(0.0401)
Prior Confidence (Std.)	0.00530	-0.00357
	(0.0288)	(0.0178)
No. of subjects	1802	1802

Note: The table reports IV regressions with the height of the submitted bids as outcome variables. Column (1) reports the results for the high and and Column (2) for the low salary positions, respectively. The risk measures are instrumented by using the ORIV approach following Gillen et al. (2019). Clustered standard errors at the Stage-1 group level in parentheses. All specifications control for field of study, occupation, the gender composition of the groups.* for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

[23] As this approach requires to duplicate each individual, the sample size doubles. Gillen et al. (2019) suggest using clustered standard errors at the participant level to receive consistent estimates of the asymptotic standard errors. To stay consistent with the main specification, we cluster standard errors at the Stage-1-group level. Following Gillen et al. (2019) by using individual level clustered standard errors does not change the standard errors much but turns the previously significant ($p = 0.088$) effect of risk on applications to the low salary position insignificant ($p = 0.112$). See Table A5 for these results.

Table 5 includes the instrumented risk as well as the confidence measures. For low salary positions, more risk loving participants submit higher bids compared to more risk averse participants. This effect is only significant at a ten percent level and is not robust to using clustered standard errors at the individual level (see Table A5 in the Appendix). Importantly, the inclusion of risk attitudes and confidence does not affect the size or significance of the main coefficients. We do not find evidence that risk attitudes or prior confidence can explain the gender gap in high salary positions.

To sum up, risk attitudes or self-confidence while being a leader are unlikely to be crucial drivers of the gender application gap for high salary positions.

5.3 Stereotypes

Leadership positions, especially those with high salaries, are primarily occupied by men (McKinsey and LeanIn.Org, 2024) and might thereby be seen as male-typed domains by society (Eagly, 1987; Badura et al., 2018; Koenig et al., 2011; Powell et al., 2002). Therefore, it is possible that the increase in financial incentive alters subjects' perceptions of the leadership role. To study this, we elicited these perceptions at the end of the experiment similar to the approach in Coffman et al. (2023): participants indicate on sliders whether the positions appear more suitable for men (1) or women (-1). As can be seen in Figure 3, participants on average state that the leadership position is rather suitable for women if there is *no* (mean = -0.023, p=0.016) or a *low* (mean = -0.016, p=0.067) payment but suitable for men if there is a *high* (mean = 0.044, p<0.01) payment. The stated suitability for the *no* as well as the *low* payment position differs from the *high* salary position at a 1% significance level, respectively. Table A6 in the Appendix shows the mean answers separately for male and female respondents. The answers by women are in line with those average results. Men's answers differ significantly from women's (see Table A7 in the Appendix) stating that all positions are more suitable for men than for women (positive means, n.s. for no salary position) but they see their strongest fit in the high salary position. In line with the findings of the non-incentivized slider questions, women believe that there are relatively more women than men among the top performers in the no and low salary position. These beliefs were elicited using incentivized belief questions[24]. They further believe that they do not differ significantly from men's performance in the high salary position. Men do not express significant differences in their beliefs about top performers.

Hence, the switch from a low to a high salary induces a change in the perception of the position from a female-typed to a male-typed task, *although* neither the task itself nor anything else, except for the salary increase, changes. This complements findingy by Babcock et al. (2017) in showing that women are more often those who perform low promotable tasks, i.e. tasks which are not helpful for evaluations and career-advancements and can be considered to be non-revenue-generating tasks

[24]These questions referred to the subject's session. Example Stage 1: "What percentage of the male team leaders in Stage 1, in which the team leaders did not receive any remuneration, managed to convince the team members to contribute their entire endowment, i.e. €3.50 per team member?" Subjects received an additional payment of 0,10€ per answer if the answer did not deviate more than 3 percentage points from the true share.

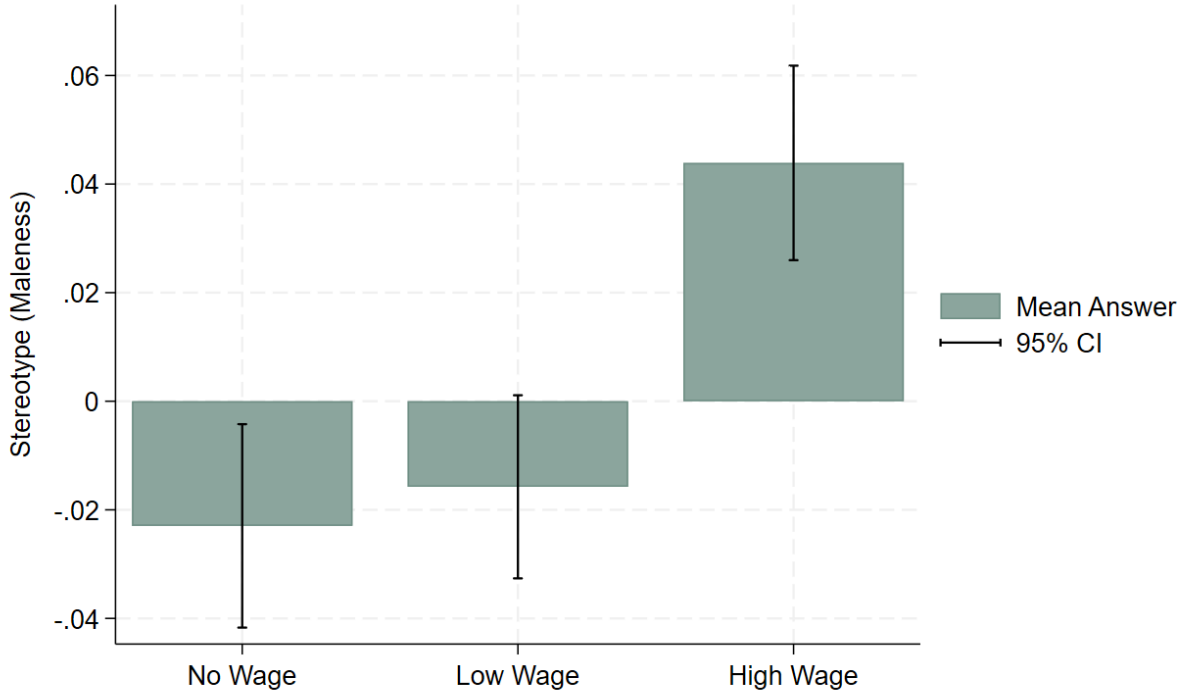

Figure 3: Stereotypical Thinking.
Note: The bars represent the mean answers to the sliders asking about the suitability of men (1) and women (-1) for the different leader positions. Confidence bars represent the 95% confidence intervals.

in industry contexts (Babcock et al., 2017). Our results show that this is already triggered by a mere change in the material benefits that can be derived from the task. Attaching a high salary induces hence a shift in stereotypes which turns the leadership position into a stereotypical male position.

To further investigate the relationship between the submitted bids and stereotypical thinking, we study how the change in bids is related to the changes in perceived gender fit for the low and high salary position. In this analysis, we calculate the change in bids (ΔBids) as the difference between the bid for the high salary position and the low salary position. Similarly, we measure the perceived fit of the respective leadership position with an individual's own gender by taking the difference between the answer on the slider: for men, we subtract the answer referring to the low from the one referring to the high salary position ($\Delta Stereot. = SliderHighS. - SliderLowS.$) so that more positive values indicate a more stereotypical thinking, in which men are believed to be relatively more suitable for the high salary position as compared to women and the low salary position. For women, we take the difference between the slider answers concerning the low and high salary position ($\Delta Stereot. = SliderLowS. - SliderHighS$) so that more positive values indicate a counter-stereotypical thinking, in which women are believed to be relatively more suitable for the high salary position as compared to men and the low salary position.

Regressing the difference in bids on the difference in stereotypes (see Table 6, Column 1) shows

Table 6: Relationship between Salary Difference and Stereotypes

	Men	Women
	(1)	(2)
	Δ Bids (Std.)	Δ Bids (Std.)
Δ Stereot. (Std.)	0.121*	-0.00366
	(0.0713)	(0.0493)
Constant	1.780***	-0.364***
	(0.353)	(0.0108)
R-squared	0.0453	0.0230
No. obs.	419	452

Note: The table reports OLS regressions with the standardized difference in bids for the high and the low salary position as the outcome variables. Clustered standard errors at the Stage 1 group level in parentheses. All specifications control for the field of study and the occupation. Column 1 shows the results for women only, Column 2 restricts the sample size to male participants. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

that for men there is indeed a positive relationship between the standardized slider measure and the standardized difference in bids: those men who hold more stereotypical beliefs about the gender fit for a position are also those who are more willing to take over the high salary position than the low salary position. This relation does not differ significantly by gender (see Column 1 in Table A8 in the Appendix) but for women, we do not find a significant relationship between their belief about the gender fit (slider) and their bidding behavior (Column 2 in Table 6)[25].

Hence, the money attached to the position comes with a change in the perceived gender congruence of the role and, especially for men, translates into subjects' willingness to pay for the highly paid positions[26].

5.4 Male-typed Positions and the Willingness to Claim the Lead

In the previous section, we have shown that subjects perceive the high salary position to be a male-typed task. It is therefore interesting to explore whether the low salary leadership positions are not perceived as a proper leader position. The low salary could contradict the common understanding of a usual wage of leaders and thereby affect the norm of who is a suitable candidate for the position.

Support for this hypothesis provide subjects' preferences to claim or grant the lead as measured by the scale of Giessner et al. (2022). A sample item for the willingness to claim the lead is "I am open to the position of leading someone" and grant the lead includes for example "I don't mind following someone's lead for what needs to get done". Table 7[27] reports the (standardized) survey

[25]Table A8 in the Appendix repeats the analysis and includes the Stage 1 and 3 leader information and an indicator for being in a low or high salary group in Stage 3. The results remain similar but become insignificant if the sample is restricted to men.

[26]Table A9 in the Appendix shows that the difference emerges because subjects bid more for the high salary position (see Columns 1 to 3) rather than less for the low salary position (see Columns 4 to 6).

[27]The sample size is reduced compared to the previous analyses because some subjects are included in the main sample whose group partners quit the experiment so that they were not able to proceed to these questions. Further, one session had technical issues so that the survey was not displayed to those 33 subjects.

Table 7: Willingness to Claim / Grant the Lead

	(1) Claim the Lead (Std.)	(2) Grant the Lead (Std.)
Female	-0.278***	0.0479
	(0.0691)	(0.0702)
Constant	1.762***	-1.211***
	(0.0691)	(0.0702)
R-squared	0.0445	0.0185
No. obs.	838	838

Note: The table reports OLS regressions with the standardized mean index of answers to the survey on claiming the lead in Column (1) ($\alpha > 0.89$) and granting the lead ($\alpha > 0.82$) in Column (2). Clustered standard errors at participant level in parentheses. All specifications control for field of study and occupation. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

answers. The answers show that men and women do not differ significantly in their willingness to grant the lead. However, there is a significant and sizable gender gap in the willingness to claim the lead: women report ceteris paribus a 0.278 standard deviations lower willingness to claim the lead than men. Hence, there indeed is a gender difference in the preference to assume leadership positions.

Besides this gender gap, Table 8 reveals that subjects who exhibit a stronger willingness to claim the lead (Column 1) and a lower willingness to grant the lead (Column 2) are also those who show larger differences in their bids for the two differently rewarded leadership positions. These relations do not significantly differ by gender but women exhibit a lower preference to claim the lead. The significant association between the preferences to claim and grant the lead with the bid differences for the two different positions suggests that indeed only the high salary position is perceived to be a true leadership position. Further supportive evidence for this argument is provided by Table A10 in the Appendix: it shows that the relation between the willingness to claim the lead and the bidding behavior is more strongly driven by the increased bids for the high salary position rather than by decreased bids for the low salary position. Also, the willingness to grant the lead is only significantly negatively related to bids for the high salary position but not for the bids for the low salary position (see Table A10 in the Appendix). Hence, especially for the high salary position the gender differences in preferences for leadership matter. This could - along with the change in stereotype - contribute to the gender application gap for high salary positions: high salaries may support the societal and internalized norms about leadership positions and leaders. These stereotypical perceptions of suitable leaders deter women from seeking those leadership roles. Men, on the other hand, are more likely to view high-salary positions as more appropriate opportunities to assert leadership than the low salary positions.

Table 8: Survey outcomes

	(1) Δ Bids (Std.)	(2) Δ Bids (Std.)
Female	-0.130*	-0.186***
	(0.0691)	(0.0689)
Claim the Lead (Std.)	0.183***	
	(0.0433)	
Female × Claim the Lead (Std.)	0.0725	
	(0.0617)	
Grant the Lead (Std.)		-0.141***
		(0.0420)
Female × Grant the Lead (Std.)		0.0629
		(0.0670)
Constant	-0.615***	-0.271***
	(0.109)	(0.0975)
R-squared	0.0726	0.0369
No. obs.	838	838

Note: The table reports OLS regressions with the standardized bid difference as outcome variables. *Claim* and *Grant* are the mean index of answers to the survey on claiming the lead in Column (1) ($\alpha > 0.89$) and granting the lead ($\alpha > 0.82$) in Column (2). Clustered standard errors at participant level in parentheses. All specifications control for field of study and occupation. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

6 Conclusion

This study investigates the willingness to pursue leadership positions, in order to better understand why women are less inclined than men to step into these roles. This experimental study examined the effect of experience and different salaries on the willingness to invest to attain leadership positions.

While the findings reveal that the exposure to leadership experience did not affect the application behavior of men and women (differently), an important distinction emerged: a gender application gap for leadership positions only emerges for positions that come with large salary increases but not for low salaries. Women exhibit a lower willingness to pay compared to men for assuming a leadership position although the leadership responsibility, task complexity, and other task dimensions remain the same. We furthermore find that the mere shift in salaries changes the perception of leadership roles from stereotypical female to male stereotyped. With this shift in perception, women's generally lower willingness to take on leadership roles (as measured by the *Claim the Lead* scale from Giessner et al. (2022)) becomes a relevant factor in their application behavior.

Does this imply that women should be paid less to bring them into leadership position? Our

results speak against this naive idea. While the increase in monetary incentive causes the gender gap to emerge, stereotypes and norms about who is deemed suitable for leadership play a critical role. These stereotypes are triggered by a mere change in salary levels that come with the position as the high salary in our study shifts gender stereotypes about the position substantially. As such, the results highlight the need for interventions that explicitly address and challenge the societal norm that equates leadership with masculinity, particularly for positions that are perceived as high-status or high-reward.

We, however, also found that the intuitively appealing policy suggestion of exposing men and women to the leadership positions to reduce the stereotypical thinking does not seem to work in our setting. Experience alone does not suffice to overcome the instilled stereotypes. Delfino (2024) has recently shown that, for instance, providing information about past performances of other employees might be a helpful tool to attract men in a female-dominated sector. It remains to be tested whether this can be adapted to applications for highly paid leadership positions.

From a practical standpoint, this study underscores the importance of rethinking how leadership roles are communicated. Emphasizing the collaborative and communicative aspects of leadership, rather than its stereotypically male-typed characteristics, could help reduce the gender gap in leadership aspirations as those are not the aspects deterring women from pursuing the leadership positions. The work of Fuchs et al. (2024) and Opitz (2025), for instance, shows that highlighting aspects of work which are appreciated in particular by women (such as working from home options or job flexibility) can increase their application rates. Stressing the communication aspect may also be used to reduce ambiguity about the job which has been shown to encourage women to enter male-dominated fields (Coffman et al., 2024). Another approach which has been discussed in the literature is to directly encourage women to enter male-dominated fields (Del Carpio and Guadalupe, 2022). It remains an empirical question whether these approaches also help to overcome stereotypes about leadership.

Future research should build on these findings to explore more effective interventions that target and reshape stereotypical thinking about leadership. Such efforts are not only critical for achieving gender balance in leadership positions but also for unlocking the potential of a more diverse talent pool, which could have profound benefits for organizations and society as a whole.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Alan, S., Ertac, S., Kubilay, E., and Loranth, G. (2020). Understanding Gender Differences in Leadership. *The Economic Journal*, 130(626):263–289.

Antonakis, J., d'Adda, G., Weber, R. A., and Zehnder, C. (2022). "just words? just speeches?" on the economic value of charismatic leadership. *Management Science*, 68(9):6355–6381.

Azmat, G., Cuñat, V., and Henry, E. (2024). Gender promotion gaps and career aspirations. *Management Science, forthcoming*, pages 1–15.

Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, 107(3):714–747.

Badura, K. L., Grijalva, E., Newman, D. A., Yan, T. T., and Jeon, G. (2018). Gender and leadership emergence: A meta-analysis and explanatory model. *Personnel Psychology*, 71(3):335–367.

Bandiera, O., Prat, A., Hansen, S., and Sadun, R. (2020). Ceo behavior and firm performance. *Journal of Political Economy*, 128(4):1325–1369.

Barker, E. H. and Gil, R. (2023). How does purpose shape the gender paygap? evidence from us charities. *SSRN Working Paper*.

Benson, A., Li, D., and Shue, K. (2024). Potential and the gender promotions gap. *Available at SSRN*.

Bircan, C., Friebel, G., and Stahl, T. (2024). Gender promotion gaps in knowledge work: The role of task assignment in teams. *Working Paper*, pages 1–55.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Bohnet, I., Van Geen, A., and Bazerman, M. (2016). When performance trumps gender bias: Joint vs. separate evaluation. *Management Science*, 62(5):1225–1234.

Cahuc, P., Carcillo, S., and Zylberberg, A. (2014). *Labor economics*. MIT press, 2 edition.

Cappelen, A. W., Reme, B.-A., Sørensen, E. Ø., and Tungodden, B. (2016). Leadership and incentives. *Management Science*, 62(7):1944–1953.

Carlsson, M. and Eriksson, S. (2019). In-group gender bias in hiring: Real-world evidence. *Economics Letters*, 185:108686.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chen, Y., Katuščák, P., and Ozdenoren, E. (2013). Why cant a woman bid more like a man? *Games and Economic Behavior*, 77(1):181–213.

Coffman, K., Collis, M. R., and Kulkarni, L. (2023). Stereotypes and belief updating. *Journal of the European Economic Association*, pages 1–44.

Coffman, K. B., Araya, P. U., and Zafar, B. (2021). A (dynamic) investigation of stereotypes, belief-updating, and behavior. *National Bureau of Economic Research Working Paper 29382*, pages 1–73.

Coffman, K. B., Collis, M. R., and Kulkarni, L. (2024). Whether to apply. *Management Science*, 70(7):4649–4669.

Day, D. V. (2000). Leadership development:: A review in context. *The Leadership Quarterly*, 11(4):581–613.

Del Carpio, L. and Guadalupe, M. (2022). More women in tech? evidence from a field experiment addressing social identity. *Management Science*, 68(5):3196–3218.

Delfino, A. (2024). Breaking gender barriers: Experimental evidence on men in pink-collar jobs. *American Economic Review*, 114(6):1816–53.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Psychology Press.

Eisenkopf, G. (2014). The impact of management incentives in intergroup contests. *European Economic Review*, 67:42–61.

Eisenkopf, G. (2020). Words and deeds–experimental evidence on leading-by-example. *The Leadership Quarterly*, 31(4):101383.

Exley, C. L. and Kessler, J. B. (2022). The gender gap in self-promotion. *The Quarterly Journal of Economics*, 137(3):1345–1381.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica*, 90(3):1063–1084.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Fluchtmann, J., Glenny, A. M., Harmon, N. A., and Maibom, J. (2024). The gender application gap: Do men and women apply for the same jobs? *American Economic Journal: Economic Policy*, 16(2):182–219.

Fuchs, L., Heinz, M., Pinger, P., and Thon, M. (2024). How to attract talents? field-experimental evidence on emphasizing flexibility and career opportunities in job advertisements. *ECONtribute Discussion Paper No. 332*, pages 1–45.

Gächter, S., Nosenzo, D., Renner, E., and Sefton, M. (2012). Who makes a good leader? cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, 50(4):953–967.

Gächter, S. and Renner, E. (2018). Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior & Organization*, 154:321–334.

Giessner, S. R., Atterstam, B., Isaakyan, S., and Jacobs, G. (2022). Willingness to claim the lead and to grant the lead – a new scale, its antecedents and outcomes. *Academy of Management Annual Meeting; Seattle, Washington, USA, August 2022*.

Gillen, B., Snowberg, E., and Yariv, L. (2019). Experimenting with measurement error: Techniques with applications to the caltech cohort study. *Journal of Political Economy*, 127(4):1826–1863.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.

Güth, W., Levati, M. V., Sutter, M., and Van Der Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, 91(5-6):1023–1042.

Haegele, I. (2024). The broken rung: Gender and the leadership gap. *Working Paper*.

Koenig, A. M., Eagly, A. H., Mitchell, A. A., and Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, 137(4):616–642.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The value of bosses. *Journal of Labor Economics*, 33(4):823–861.

Lecocq, S., Magnac, T., Pichery, M.-C., and Visser, M. (2005). The impact of information on wine auction prices: results of an experiment. *Annales d'Economie et de Statistique*, pages 37–57.

Martin, R., Hughes, D. J., Epitropaki, O., and Thomas, G. (2021). In pursuit of causality in leadership training research: A review and pragmatic recommendations. *The Leadership Quarterly*, 32(5):101375.

McKinsey and LeanIn.Org (2023). Women in the workplace 2023.

McKinsey and LeanIn.Org (2024). Women in the workplace 2024.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Onderstal, S. (2020). Premium auctions in the field. *Review of Economic Design*, 24(1):39–63.

Opitz, S. (2025). Highlighting benefits in job ads: A field experiment. *Unpublished Manuscript*.

Powell, G. N., Butterfield, D. A., and Parent, J. D. (2002). Gender and managerial stereotypes: have the times changed? *Journal of Management*, 28(2):177–193.

Rutström, E. E. (1998). Home-grown values and incentive compatible auction design. *International Journal of Game Theory*, 27:427–441.

Smith, N. (2018). Gender quotas on boards of directors. *IZA World of Labor*, 7.

SpencerStuart (2024). 2024 s&p 500 board diversity snapshot.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37.

Weber, R., Camerer, C., Rottenstreich, Y., and Knez, M. (2001). The illusion of leadership: Misattribution of cause in coordination games. *Organization Science*, 12(5):582–598.

A Appendix

Figure A1: Distribution of Bids

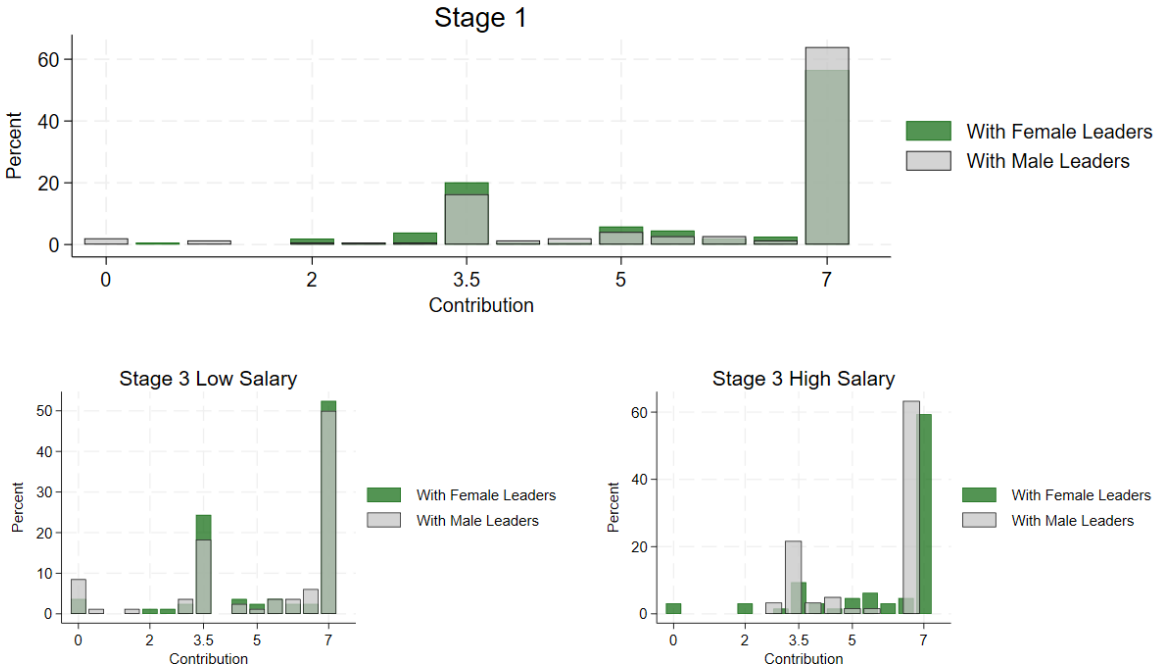

Figure A2: Team Members' contributions in all stages

Table A1: Balance table

Variable	(1) Pooled Sample	(2) Team Members	(3) Team Leaders	(4) Difference (3)-(2)
Female	0.514	0.515	0.512	-0.003
	(0.500)	(0.500)	(0.501)	(0.035)
Age	27.920	28.095	27.571	-0.524
	(7.385)	(7.521)	(7.105)	(0.507)
Student	0.710	0.707	0.718	0.011
	(0.454)	(0.456)	(0.451)	(0.031)
Employee	0.166	0.168	0.163	-0.006
	(0.373)	(0.374)	(0.370)	(0.024)
Civil Servant	0.073	0.068	0.083	0.015
	(0.261)	(0.253)	(0.276)	(0.018)
Trainee	0.007	0.007	0.007	-0.000
	(0.081)	(0.081)	(0.081)	(0.006)
Self-Employed	0.019	0.018	0.020	0.002
	(0.136)	(0.134)	(0.140)	(0.010)
Unemployed	0.019	0.027	0.003	-0.023***
	(0.136)	(0.161)	(0.058)	(0.007)
Retired	0.003	0.003	0.003	-0.000
	(0.058)	(0.058)	(0.058)	(0.004)
No Answer	0.002	0.002	0.003	0.002
(Occupation)	(0.047)	(0.041)	(0.058)	(0.004)
Human Sciences	0.085	0.072	0.113	0.041*
	(0.280)	(0.258)	(0.317)	(0.021)
Mathematics &	0.084	0.097	0.060	-0.037*
Natural Sciences	(0.278)	(0.296)	(0.238)	(0.019)
Medicine	0.039	0.047	0.023	-0.023**
	(0.193)	(0.211)	(0.151)	(0.012)
Arts and Humanities	0.069	0.070	0.066	-0.004
	(0.253)	(0.255)	(0.249)	(0.017)
Law	0.054	0.065	0.033	-0.032**
	(0.227)	(0.247)	(0.180)	(0.015)
Management, Economics	0.305	0.285	0.346	0.061*
& Social Sciences	(0.461)	(0.452)	(0.476)	(0.033)
Other Faculty	0.095	0.100	0.086	-0.014
	(0.294)	(0.300)	(0.281)	(0.021)
Observations	901	600	301	901

Note: Values in parentheses indicate standard deviations (Columns (1)-(3)) and standard errors (Column (4)). Standard errors are clustered at the group level. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A2: Gender Application Gap (individual clustered standard errors)

	(1) Bid	(2) Bid High S.	(3) Bid Low S.
Female	-0.00293	-0.179**	0.0113
	(0.0523)	(0.0727)	(0.0512)
High Salary	0.863***		
	(0.0339)		
Female × High Salary	-0.162***		
	(0.0509)		
Leader Experience	0.0333	0.0559	0.0108
	(0.0501)	(0.0670)	(0.0453)
Constant	0.595**	1.470***	0.583*
	(0.282)	(0.282)	(0.306)
R-squared	0.204	0.0274	0.0149
No. obs.	1802	901	901

Note: The table reports OLS regressions with the height of the submitted bids as outcome variables. Column (1) reports the pooled results with clustered standard errors at the individual level in parentheses. Columns (2) and (3) split the sample into bids for the high and low salary positions, respectively and show robust standard errors in parentheses. All specifications control for field of study, occupation, the gender composition of the groups. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A3: Gender Application Gap

	(1) Bid	(2) Bid High S.	(3) Bid Low S.	(4) Bid	(5) Bid High S.	(6) Bid Low S.
Female	0.0208	-0.141*	0.00990	-0.00347	-0.192**	0.00230
	(0.0437)	(0.0775)	(0.0520)	(0.0529)	(0.0849)	(0.0603)
High Salary	0.863***			0.863***		
	(0.0356)			(0.0358)		
Female ×	-0.162***			-0.162***		
High Salary	(0.0514)			(0.0516)		
Leader Exp.	0.0391	0.0583	0.00338	0.0330	0.0415	-0.00727
	(0.0502)	(0.0853)	(0.0559)	(0.0516)	(0.0878)	(0.0566)
Female ×		-0.000370	0.0324		0.0216	0.0407
Leader Exp.		(0.132)	(0.0829)		(0.135)	(0.0850)
Constant	0.376***	1.233***	0.388***	0.619**	1.739***	0.358
	(0.0342)	(0.0514)	(0.0367)	(0.301)	(0.323)	(0.315)
Controls	No	No	No	Yes	Yes	Yes
R-squared	0.192	0.00634	0.000627	0.204	0.0306	0.0197
No. obs.	1802	901	901	1802	901	901

Note: The table reports OLS regressions with the height of the submitted bids as outcome variables. Column (1) & (4) report the pooled results, Columns (2),(3), (5), and (6) split the sample into bids for the high and low salary positions. Clustered standard errors in parentheses. Controls: field of study, occupation, the gender composition of the groups, age. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A4: Gender Gaps in the Extensive Margin

	(1) Apply	(2) Apply High S.	(3) Apply Low S.
Female	-0.133***	-0.0963***	-0.130***
	(0.0366)	(0.0313)	(0.0391)
High Salary	0.167***		
	(0.0182)		
Female × High Salary	0.0385		
	(0.0267)		
Leader Experience	0.0517*	0.0568*	0.0466
	(0.0280)	(0.0289)	(0.0347)
Constant	0.736***	1.034***	0.605*
	(0.168)	(0.0853)	(0.316)
R-squared	0.0736	0.0303	0.0463
No. obs.	1802	901	901

Note: The table reports OLS regressions with a dummy for non-zero bids as outcome variables. Column (1) reports the pooled results, Columns (2) and (3) split the sample into bids for the high and low salary positions, respectively. Clustered standard errors in parentheses. All specifications control for field of study, occupation, the gender composition of the groups. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A5: Risk and Confidence

	(1) Bid High S.	(2) Bid Low S.
Female	-0.175**	0.0152
	(0.0720)	(0.0506)
Leader Experience	0.0569	0.0116
	(0.0662)	(0.0447)
Risk (Instr.)	0.0520	0.0686
	(0.0575)	(0.0431)
Prior Confidence (Std.)	0.00530	-0.00357
	(0.0300)	(0.0181)
No. of subjects	1802	1802

Note: The table reports IV regressions with the submitted bids as outcome variables. Column (1) reports the results for the high and and Column (2) for the low salary positions, respectively. The risk measures are instrumented with the ORIV approach following Gillen et al. (2019). Clustered standard errors at the individual level in parentheses. All specifications control for field of study, occupation, the gender composition of the groups.* for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Position	Average Response by Men	Average Response by Women	Overall Average
No Salary	0.016 (p=0.167)	-0.059*** (p < 0.01)	-0.023* (p = 0.016)
Low Salary	0.023** (p=0.047)	-0.052*** (p<0.01)	-0.016* (p = 0.067)
High Salary	0.078*** (p<0.01)	0.012 (p=0.351)	0.044*** (p<0.01)

Table A6: Perceived Gender-Fit

Notes: The table reports the average responses to the slider questions asking about the suitability of men and women to be a leader with the three different salary options by gender. P-values refer to t-tests against zero. * for p < 0.10, ** for p < 0.05 and *** for p < 0.01

Table A7: Gender Differences Perceived Gender Fit

	(1) Maleness No Salary	(2) Maleness Low Salary	(3) Maleness High Salary
Female	-0.0755***	-0.0685***	-0.0612***
	(0.0193)	(0.0171)	(0.0182)
Constant	0.0755***	0.0685***	0.0612***
	(0.0193)	(0.0171)	(0.0182)
R-squared	0.0337	0.0345	0.0235
No. obs.	871	871	871

Note: The table reports OLS regressions with the answer to the slider questions as outcome variables ranging from -1 (suitable for women) to 1 (suitable for men). Robust standard errors in parentheses. All specifications control for the field of study and the occupation. * for p < 0.10, ** for p < 0.05 and *** for p < 0.01.

Table A8: Relationship bid differences and stereotypes

	All	Men	Women
	(1)	(2)	(3)
	Δ Bids (Std.)	Δ Bids (Std.)	Δ Bids (Std.)
Δ Stereot. (Std.)	0.121*	0.118	0.0191
	(0.0726)	(0.0721)	(0.0485)
Female	-0.274***		
	(0.0877)		
Female × Δ Stereot. (Std.)	-0.108		
	(0.0865)		
Constant	-0.371	1.594***	-0.793***
	(0.248)	(0.410)	(0.235)
R-squared	0.0868	0.0808	0.0963
No. obs.	871	419	452

Note: The table reports OLS regressions with the standardized difference in bids for the high and the low salary positions as the outcome variables. Clustered standard errors at participant level in parentheses. All specifications control for the field of study, the occupation, the gender compositions of groups in Stages 1 and 3, being a leader in Stage 1 or 3 and being in the low or high salary group. Column 2 shows the results for men only, Column 3 restricts the sample size to female participants. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A9: Relationship actual bids and stereotypes

	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
	Bid High S.	Bid High S.	Bid High S.	Bid Low S.	Bid Low S.	Bid Low S.
Δ Stereot.	0.0601*	0.0585*	-0.0680	-0.0337	-0.0343	-0.0652
(Std.)	(0.0346)	(0.0353)	(0.0598)	(0.0383)	(0.0388)	(0.0479)
Female	-0.155**			0.0116		
	(0.0631)			(0.0445)		
Female ×	-0.131*			-0.0318		
Δ Stereot. (Std.)	(0.0671)			(0.0608)		
Constant	1.670***	2.638***	1.515***	1.003***	0.492	1.014***
	(0.0648)	(0.631)	(0.0131)	(0.0436)	(0.360)	(0.0105)
R-squared	0.0283	0.0376	0.0322	0.0212	0.0349	0.0307
No. obs.	871	419	452	871	419	452

Note: The table reports OLS regressions with the submitted bids for the high (Columns 1-3) and the low salary position (Columns 4-6) as the outcome variables. Clustered standard errors at the Stage 1 group level in parentheses. All specifications control for the field of study and the occupation. Columns 2 and 5 show the results for men only, Columns 3 and 6 restrict the sample size to female participants. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A10: Survey outcomes separate bids

	All	Women	Men	All	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)
		Bid High S.			Bid Low S.	
Claim the Lead	0.274***	0.372***	0.174***	0.113***	0.177***	0.0531
(Std.)	(0.0334)	(0.0459)	(0.0503)	(0.0250)	(0.0364)	(0.0357)
Grant the Lead	-0.0793**	-0.0203	-0.129***	-0.0242	-0.00239	-0.0331
(Std.)	(0.0324)	(0.0496)	(0.0419)	(0.0227)	(0.0347)	(0.0285)
Female	-0.0480			0.0510		
	(0.0645)			(0.0460)		
Constant	1.049***	0.925***	2.538***	0.753***	0.734***	0.475
	(0.0958)	(0.0778)	(0.645)	(0.0674)	(0.0562)	(0.364)
R-squared	0.117	0.164	0.0993	0.0489	0.0905	0.0456
No. obs.	838	429	409	838	429	409

Note: The table reports OLS regressions with the bids for the high salary positions in Columns (1) - (3) and the low salary positions in Columns (4) - (6). *Claim* and *Grant* are the mean index of answers to the survey on claiming the lead and granting the lead. Clustered standard errors at participant level in parentheses. All specifications control for field of study and occupation. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.

Table A11: Leaders' Payments Stage 3

	No Salary	Low Salary	High Salary	No Salary	Low Salary	High Salary
	(1)	(2)	(3)	(4)	(5)	(6)
		Total Group Contribution			Leaders' Earnings	
Female	-0.175	-0.182	-0.227	0.0843	0.273	0.295
	(0.112)	(0.166)	(0.182)	(0.267)	(0.431)	(0.431)
Constant	5.782***	3.500***	5.983***	10.48***	7.971***	11.07***
	(0.961)	(0.648)	(1.154)	(0.473)	(1.681)	(2.736)
R-squared	0.0560	0.143	0.0876	0.0210	0.101	0.110
No. obs.	301	124	164	301	124	164

Note: The table reports OLS regressions with the groups' overall contributions as outcome variables in Columns (1)-(3) and team leaders' payoffs (without salary) as outcome variables in Columns (4) - (6). Columns (1) and (4) refer to Stage 1 leaders, Columns (2) and (5) to leaders in the high salary position and Columns (3) and (6) to leaders in the low salary position. All regressions control for the gender of team members, field of study and occupation. Robust standard errors in parentheses. * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$.