Department of Economics
ISSN number 1441-5429
Discussion number 14/20

The Gender Leadership Gap: Insights from Experiments

Catherine Eckel[1], Lata Gangadharan[2], Philip J. Grossman[3], Nina Xue[4]

1.Eckel: Texas A&M University, College Station, Texas, United States of America. ceckel@tamu.edu	2.Gangadharan: Monash University, Clayton, Victoria, Australia. Lata.Gangadharan@monash.edu.
3.Grossman: Monash University, Clayton, Victoria, Australia. Philip.Grossman@monash.edu.	4.Xue: Monash University, Clayton, Victoria, Australia. Nina.Xue@monash.edu.

Corresponding author: Philip Grossman
Department of Economics
Monash University
Wellington Road
Clayton
Victoria 3800
Australia
Telephone: + 61-3-9903-1400
E-mail: Philip.Grossman@monash.edu

Forthcoming in: **A Research Agenda for Experimental Economics,** edited by A. Chaudhuri. Edward Elgar.

Acknowledgement: The authors gratefully acknowledge research assistance provided by Evelina Onopriyenko.

© @2020 **Catherine Eckel, Lata Gangadharan, Philip J. Grossman and Nina Xue**
All rights reserved. No part of this paper may be reproduced in any form, or stored in a retrieval system, without the prior written permission of the author.

monash.edu/businesseconomics
ABN 12 377 614 012 CRICOS Provider No. 00008C

1. Introduction

Differences in the economic decisions and labor market outcomes of women and men are widely documented. Women have lower earnings, work in different occupations, and invest in lower risk retirement portfolios. They are less likely to be found in leadership positions, and, when leading, may be evaluated and rewarded differently.

Women's underrepresentation in leadership positions may arise from several different sources. First, women might choose differently (i.e., women may be less likely to prefer a leadership role). Second, women may be less likely to be selected for a leadership position. Third, women's effectiveness may be perceived and evaluated differently from men's, jeopardizing their success as leaders.

Considerable evidence documents gender differences in preferences and attitudes, showing that women are on average more risk averse and altruistic, and less competitive and confident than men (Niederle, 2015).[i] In turn, many studies show that preferences impact labor market outcomes; for example, more risk averse individuals select professions with lower earnings risk as in the public sector, or more altruistic individuals select low-paying care occupations such as education or health care (Shurchkov and Eckel, 2018). Evidence also indicates that women are reluctant to volunteer for or accept a leadership position, even among those who are at least as competitive and risk tolerant as their male counterparts. These differences in preferences or attitudes might explain why women, on average, would be less willing to choose to become leaders.

Women may also be less likely to be selected as leaders. This could be due to stereotypes, beliefs about preferences or ability, or discrimination. For example, opportunities for women to be hired into risky positions could be constrained by employers' beliefs, effectively barring women from some professions or leadership positions. Traditional gender roles may leave potential employers with a belief that women are simply not suitable for leadership.

Finally, leaders do not exist in a vacuum: Social interactions and norms affect perceptions of leader performance. The evaluations of leaders may also be shaped by gender-based stereotypes

and beliefs. It is often said that a woman has to be twice as good at her job as a man to be seen as equally competent. Social interactions, both with decision makers in the leader selection process and with potential followers, are therefore important to the success of a leader. Environments that are male-dominated or generally considered to be masculine (such as STEM fields) are especially susceptible to stereotypes that fuel negative reactions to female candidates and contribute to women's reluctance to assume leadership roles.

These factors may explain the underrepresentation of women in leadership roles, as well as the mixed success of interventions and policies that aim to correct for historical inequalities by providing more opportunities for women.

Most of the research on gender and the labor market relies on observational data to document and explore underlying causes of gender differences (Blau and Kahn, 2017). But it can be difficult or impossible to disentangle potential causal relationships or to obtain appropriate counterfactuals using these data. Looking only at the outcomes of the decision process, it is impossible to identify the causal effects of preferences, or to separate preferences from the effects of stereotypes or discrimination. Research using experimental methods provides a way to examine gender differences, isolating preferences from constraints. Preferences can be elicited one at a time, and ways in which women are treated can be observed without confounds such as performance and candidate quality.

Experimental research is useful in testing potential remedies for differences in outcomes. Lab experiments provide a low-cost "wind tunnel" option for pretesting responses to specific interventions. Both lab and field experiments have been designed to test the effectiveness of interventions aimed at improving gender equality. Interventions may attempt to change preferences by making women more risk tolerant or competitive, or may implement new institutions such as hiring or personnel evaluation practices. Experiments also can be designed to shed light on the mechanisms underlying an intervention's success or failure.

In this chapter, we offer a review of experimental research that we see as relevant for understanding the gender gap in leadership. Our primary focus is on lab experiments and lab-in-the-field experiments (Eckel and Candelo, 2020).[ii] This research area is large and growing, and our space is limited; necessarily, our review leaves out many important and interesting papers. Some of the papers we include examine factors that affect differences in labor market outcomes more generally, while others focus more closely on leadership, per se. Our purpose is to highlight studies that help us understand the gender gap in leadership, and the potential effectiveness of measures that might close that gap.

The chapter proceeds as follows. Section 2 focuses on women's willingness to lead. We address the gender difference in preferences, summarizing the most common experimental findings and their relevance for the gender leadership gap. Section 3 examines the selection of women as leaders, and reviews work that explores differences in perceptions, beliefs, and behavior that may contribute to the leadership gap. We discuss stereotypes, which play an important role, not just in this section, but in all aspects of the gender leadership gap. We also discuss evidence of discrimination in the selection of leaders. Section 4 discusses the ways in which male and female leaders may differ in how they are evaluated as well as in their effectiveness as leaders. The focus of Section 5 is on interventions and their effectiveness in terms of improving outcomes for women. The final section discusses policy directions (e.g., affirmative action and leader selection methods) and future research.

2. Women's decision to lead

Not everyone has what it takes to be a leader. The role of leader requires specific talents and a level of confidence in one's ability. Even if a person has the required talents and confidence, the desire to take on the role may be lacking. The process of becoming a leader is often competitive, and leaders are expected to take risks for their group or organization, to take responsibility for their own decisions and their group's success, and to be willing to sacrifice for the team when needed. Not everyone wishes to put themselves under this strain. As such, preferences are likely to influence

both the decision to pursue leadership, as well as the decisions made as leaders. In this section, we review the research on gender differences in relevant preferences, confidence, and the desire to lead.

2.1 Preferences: risk aversion, competitiveness, altruism

Research on gender differences in preferences in a leadership context has largely focused on differences in risk aversion, competitiveness and other-regarding preferences, including altruism, inequality aversion, and cooperation (Croson and Gneezy, 2009; Niederle 2015).

Incentivized and survey risk-elicitation measures find consistent evidence that women are more risk averse, though the magnitude of the difference varies depending on the method and target population (Byrnes et al., 1999; Eckel and Grossman, 2008).[iii] Greater risk tolerance is associated with the choice of STEM fields by high school students (Buser et al., 2017), selection of riskier investment portfolios (Jianakoplos and Bernasek 1998), and the choice of riskier occupations (Bonin et al., 2007). Even among CEOs, women select less risky compensation packages, with a smaller proportion of their incomes based on firm performance (Chauvin and Ash, 1994). In the lab, women choose less risky real-effort tasks to determine their earnings, with a resulting loss in earnings (Jung et al., 2018).

Entrepreneurship is another manifestation of leadership that is related to preferences. Jiang and Capra (2018) find women are less likely to be entrepreneurs, but do not observe a connection between risk tolerance and entrepreneurship.[iv] However, in a sample of small business managers in Peru, Castillo et al. (2010) find that entrepreneurs have greater risk tolerance than hired managers. Since women are on average more risk averse, this may help explain their underrepresentation among entrepreneurs.

A preference for competition has implications for the decision to lead. Rising through the ranks of an organization is a competitive process not unlike participating in a winner-take-all tournament. To assess gender differences in competitiveness, Niederle and Vesterlund (2007) developed a multistage experimental paradigm that first asks subjects to complete a task for a piece rate, then complete the same task under a tournament (competitive) setting, and finally choose

which they prefer for a third round of the task. This study has been replicated numerous times with different populations, and most support the finding that women are less competitive (Niederle, 2015; Niederle, 2017).[v] A preference for competition is associated with choice of occupation and earnings, and like risk tolerance, also predicts selection of math-intensive education tracks (Buser et al., 2014, 2017). For MBA students, differences in competitiveness are associated with industry choice and starting salaries after graduation (Reuben et al., 2015).

The relationship between cooperation and altruism and the decision to become a leader has at least two facets. On the one hand, we may not typically associate cooperation and altruism with leaders. On the other hand, a good leader puts the welfare of the group ahead of personal welfare. Leadership can involve taking on a task for the benefit of a group, such as chairing a committee or accepting a departmental administrative position, and altruism might play an important part in this kind of leadership role. Babcock et al. (2017) show that women are more likely to volunteer for thankless tasks (i.e., women are more willing to "take one for the team" by performing a sacrifice that benefits the other members of the group). Their result can be reinterpreted as evidence that women are more likely to volunteer to take on a particular type of leadership role. While this argument was not explicitly tested in their study, arguably, this kind of sacrifice is often expected of leaders in a group.[vi]

Women's greater altruism may contribute to their career choices and to their lower earnings.[vii] With respect to other-regarding preferences, as measured in the simple setting of a dictator game, women are typically more generous, giving more to an anonymous counterpart than men; and the difference is even more pronounced when the recipient is a needy individual or a charitable organization (Bilén et al., 2020, provide a meta-analysis). In this game, women are also observed to be more inequality averse and men more concerned with maximizing efficiency (Andreoni and Vesterlund, 2001). Altruistic decisions in the dictator game have been shown to predict the choice of public sector occupations (Tonin and Vlassopoulos, 2015; Banuri and Keefer,

2016). Altruism is also higher for workers in "care work" (education, health care) sectors (England et al., 2002; Folbre, ed., 2012).

2.2 Confidence and the decision to lead

The literature on gender and leadership generally shows that women are less likely to want to take on a leadership role in many contexts, including investment (Ertac and Gurdal 2012), public goods (Arbak and Villeval 2013), politics (Preece and Stoddard, 2015; Kanthak and Woon, 2015), or in a competition (Reuben et al., 2012; Erkal et al., 2019). While this may be partially explained by preferences, that may not be the full story. Even among women with a preference for leadership roles, they may lack confidence in their ability to be successful leaders.

Women may be less likely to consider themselves qualified for a leadership position. In male-typed domains, women are shown to be more reluctant to contribute ideas to group discussions (Bordalo et al., 2019), indicating a lack of confidence to put themselves forward. Gender stereotypes can influence women's self-assessments of how qualified they are to lead in particular fields (Coffman et al., 2019). Coffman (2014) argues that an individual's decision to contribute ideas is a function of ability, gender, and the gender stereotype associated with the domain. If a domain is stereotypically outside their gender's domain, the individual is less willing to contribute. The author's design allows her to infer that this behavior reflects self-assessment differences, not a fear of discrimination.

This gender confidence gap is also observed by Niederle and Vesterlund (2007), who assess confidence as a possible reason for women's lower competitiveness by asking participants to make incentivized guesses about their own performance in the tournament relative to others. While both genders are overconfident regarding their abilities, men are more overconfident than women. Controlling for actual performance, women are significantly less confident than men about their relative ranks. Beyer (1990) reports an overestimation of ability by men on masculine and neutral tasks while women are more likely to underestimate their performances in masculine tasks.[viii]

Exley and Kessler (2019) show that gender differences in self-assessment drive a substantial gender gap in self-promotion. Even though women slightly outperform men in a commonly-used test of ability (the Armed Services Vocational Aptitude Battery), their self-assessments show the opposite pattern: Women report less favorable evaluations of their own performance despite performing slightly better than men. This gap in self-promotion persists even after accurate information is provided about absolute and relative past performance. This difference in the willingness to self-promote offers a key insight into the underrepresentation of women in leadership positions, as many leadership selection processes rely on a candidate's self-evaluation of performance and ability.[ix]

Gender stereotypes, while applicable to the perceptions and evaluations of others may also be applied to own performance. Inzlicht and Ben-Zeev (2000) observe that women's performance in a math test declines with the number of men in their group. Since mathematics is typically considered to be a male domain, they propose that the presence of male team members strengthens the salience of the stereotype, thus creating a 'stereotype threat' (Aronson and Steele, 1995). While stereotypes may offer an exaggeration of actual differences, stereotypes themselves may be strengthened by environmental factors such as a group's gender composition.

Kamas and Preston (2018) measure the taste for competition and confidence of college seniors and then track their labor market experiences in the early years following graduation. They find that a preference for competition and confidence are positively correlated with compensation for women, but not for men. Competitive and confident women earn more than other women and just as much as men.

2.3 Willingness to accept the role of leader

A consequence of the differences in preferences and the confidence gap is seen in the reluctance of women even to accept the role of leader. Grossman et al. (2015) and Li et al. (2020) find that women are less willing to lead, particularly in situations in which the leader's gender is revealed to group members, with female leaders most willing to lead in all female groups. Similarly, Born et al. (2019)

find that women's willingness to lead is greatest in female majority groups. Subjects participate in two framed survival tasks, both individually and in a group of four (either three men/one woman or one man/three women). Individuals indicate their willingness to lead their groups (i.e., make the final decisions based on members' suggestions). Men are more willing to be the leader and are even more willing to lead in female majority groups. Men are more confident, have more influence, and are ranked higher than equally performing women.

Chen and Houser (2019) manipulate the salience of gender stereotypes to test whether a group's gender mix affects the willingness to become a leader. Participants are randomly assigned to groups with varying gender compositions, and are then asked to report their willingness to answer general knowledge questions for their groups. Women in same-gender groups are twice as willing to accept a leadership role as those in mixed-gender groups and the result is explained by a change in confidence levels. This suggests that women's confidence is particularly likely to suffer in male-dominated fields.

Ertac and Gurdal (2012) show that women are less willing to lead in a simple risky environment, deciding how much of a fixed endowment to invest in a risky asset. When placed in groups of five, women are more likely to refuse to make a decision on behalf of their group.

Ertac et al. (2019) assess how much individuals are willing to pay either to obtain the authority to make a risky decision for themselves, or to avoid the responsibility of making the decision by deferring it to the leader. Subjects are randomly allocated to the decision-maker role, but then can pay to change the situation. The experiment elicits the leader's willingness to pay to avoid the leadership position (demand to avoid responsibility, DAR), and the follower's willingness to pay to make their own decision (demand for autonomy, DFA). Someone who prefers to be a leader has negative DAR and positive DFA. Women are substantially less likely to fall into this categorization than men (25.5% of women and 43.3% of men). This study also shows a negative relationship between other-regarding preferences and the preference to be a leader.

Chakraborty and Serra (2019) design an experiment that mimics corporate decision making, where the leader faces the possibility of angry messages from subordinates. They show that "backlash aversion" can play an important role in women's decision not to pursue a leadership role.

Propp (1995) argues that a reluctance to take on leadership roles may also be due to an inability to exert influence. Attempts by women to exert influence are often resisted or ignored, men have more influence than women in mixed groups, and the contributions of men, relative to the same contribution by women, receive more attention from other group members (Carli, 2001; Gangadharan et al., 2016). We discuss the evaluation of male and female leaders further in Section 4.

Leaders are often elected, or chosen through some other process. Women's under-representation among candidates for political office is often decried in the popular press. Kanthak and Woon (2015) investigate why women don't run for office, and find that women are averse to elections. They rule out differences in preferences or confidence as reasons why women do not run, and suggest another possibility. When elections are completely truthful – that is, no one has the scope to make unwarranted claims about their qualifications – women are as likely as men to put themselves forward. Women are averse to competing in an environment in which qualifications are not verifiable, and in which exaggeration of candidates' own qualifications and false doubt about others' claims play a role.

In sum, gender differences in risk-taking, aversion to competition, altruism, confidence in own ability, self-promotion, and willingness to lead offer possible explanations for the gender gap in leadership, but the underrepresentation of women in leadership roles is not fully explained by these factors.

3. Perceptions, beliefs, and the selection of leaders

In this section, we consider the beliefs and perceptions of those involved in appointing leaders, and those of potential followers regarding female leaders, and how this affects the selection of women as leaders.

Even if women possess the same preferences, attitudes, qualities, and talents as their male counterparts, gender-based perceptions or beliefs can persist. If women are perceived to be less risk tolerant and/or less competitive, this could influence their selection for a competitive task. Indeed, there is evidence that, to be perceived as effective, women may be held to different standards than men, and may need to show that they are simultaneously sensitive (other-regarding) as well as strong (risk-taking and competitive) (Johnson et al., 2008).

3.1 Gender Stereotypes

Gender stereotypes about preferences can play an important part in leader selection. Considerable evidence supports the presence of stereotypes that exaggerate underlying true differences. For example, Eckel and Grossman (2002) show that women are stereotyped as more risk averse than men; beliefs about gender differences are greater than the actual gender differences in risk aversion. Similarly, Cason et al. (2020) find that men and women both believe that women are more community-minded and more likely to make prosocial choices and inflate the underlying difference.

Another area in which stereotypes matter is confidence. Confidence affects women's willingness to become leaders (discussed in Section 2.2), but stereotypes about confidence also are likely to affect women's selection as leaders. Bordalo et al. (2019) argue that men's overconfidence may be based in part on the exaggeration of true underlying differences, and in part on an overestimation of own ability. They confirm that stereotypes exaggerate the performance gap. Both men and women are influenced by stereotypes: In male- (female-) oriented domains, both men and women underestimate (overestimate) the ability of women relative to men.

3.2 Discrimination

Discrimination against women in the selection of leaders can be based on differences in beliefs about women and men, including gender stereotypes and statistical discrimination, or can be taste based. While the evidence of gender stereotyping is fairly conclusive, its impact on the selection of female leaders is discussed in this section.

In general, considerable evidence documents gender differences in the selection of workers. In labor markets, women are less likely to be hired into male-dominated domains, and women's credentials are often discounted.[x] Blau and Kahn (2017) survey the evidence on gender discrimination in hiring, including the pathbreaking study by Goldin and Rouse (2000) showing that implementing blind auditions for orchestras substantially increased the hiring of women.

Gender discrimination can also be seen in studies focusing on the selection of leaders. For example, Moss-Racusin et al. (2012) submitted identical applications, differing only in the gender of the applicant, for a lab manager position. Both male and female employers exhibited a preference for male applicants for this leadership position; women with identical qualifications were perceived as less qualified than male candidates.

Coffman et al. (2020) focus on distinguishing between belief and taste-based gender discrimination. They conduct an experiment in which employers select workers to complete tasks in male-stereotyped domains (math and sports). The researchers pretest their subjects' performance in a math and a sports quiz. Employers then select among available workers. In one treatment, workers are identified by gender; the control condition is cleverly designed by allocating the same set of workers into two groups depending on whether someone was born in an odd or even birth month (all male workers and all female workers, respectively), a categorization about which there are no differences in beliefs. Their design isolates the gender-specific, taste-based component of discrimination. Their clever design includes women born in even months and men born in odd months. When the labelling is changed from gender to birth month, but the same information is given about the groups, employers discriminate against even-month workers. The design creates the same set of beliefs in the two treatments, and similar behavior is observed in both. This allows them to conclude that the discrimination is belief-based, and not taste-based.[xi]

3.3 Negotiation

Even if women are willing to take on leadership roles, a further roadblock lies in the negotiation process. The process of becoming a leader often involves some form of negotiation (such as

negotiating for a promotion). Hernandez-Arenaz and Iriberri (2019a) provide an extensive survey of the literature on gender differences in negotiation.

There is substantial evidence that women are more reluctant to enter into negotiations than men (Babcock and Laschever, 2009; Eriksson and Sandberg, 2012) and this has been studied extensively in the context of wage negotiations (Dittrich et al., 2014). Gender stereotypes can be especially harmful to women during the negotiation process. When women do enter negotiations, Amanatullah and Morris (2010) find that women take a less aggressive approach in anticipation of backlash for violating gender stereotypes. In a series of experiments, Bowles et al. (2007) examine willingness to negotiate as well as the treatment of men and women who attempt to negotiate. Female candidates who initiate wage negotiations are penalized more than male candidates, and this effect appears to be driven by male evaluators.

Eckel et al. (2008) survey negotiation experiments in the simple environments of dictator and ultimatum games. They note that differences in negotiation outcomes can result from differences in the behavior of women and men, or differences in how women and men are perceived and treated. While women have a stronger preference for fair outcomes than men, this difference in behavior is small compared to the difference in expectations about how men and women behave. They also offer evidence that women may be more sensitive to the negotiation context. Small et al. (2007) report a gender difference due to the bargaining context. If the interaction is framed as a negotiation for higher earnings, women are less willing to negotiate. Reframing the interaction as an opportunity to "ask" for higher pay appears to eliminate gender differences.[xii]

On balance, stereotypes have an important and persistent influence on perceptions of others' preferences and abilities. Stereotypes can also lead to discrimination and gender differences in negotiation outcomes contributing to the leadership gap.

4. Leader evaluation and effectiveness

Gender stereotypes play a key role in the selection of female leaders, but also in the assessment of leaders and their effectiveness. Not only do these factors contribute to an ability to lead effectively,

and to be seen as such, they can also bias the negotiation process (for wages or promotions), which, as discussed in Section 3.3, is often biased against women. In this section we discuss research on the evaluation of leaders and leader effectiveness.

4.1 Gender and the evaluation of leaders

Gender stereotypes can bias the evaluation of women as leaders. This may be especially detrimental to female leaders when qualities that are desirable for a leader (such as assertiveness) are incongruent with stereotypes about women. Observationally, women appear to be evaluated differently in many domains, from teaching evaluations (Boring, 2017) to candidates running for office (Murray, 2008). We discuss experiments that test whether male and female leaders are indeed evaluated differently.

Grossman et al. (2019) employ a repeated weakest-link coordination game to address gender differences in leader effectiveness and followers' perceptions of leaders' effectiveness. The leader's intervention is a short, semiscripted speech advising followers on how to maximize earnings. The advantage of their design is that there can be no difference in leader ability or quality, removing these as explanations for variations in evaluations. Followers evaluate their leaders' effectiveness and may reward their leaders with a bonus. While followers are influenced equally by the advice of male and female leaders, female leaders are assessed less positively and rewarded less generously than equally-effective male leaders. Using a similar design, Li et al. (2020, discussed in Section 2.3) also observe no difference in followers' responses to the messages of male and female leaders.

Similar to Grossman et al, (2019), Brooks et al. (2014) find that both professional investors and nonprofessional evaluators prefer entrepreneurial pitches presented by male as compared to female entrepreneurs, even when the content of the pitch is the same. Using tenure decisions of academic economists and evidence from experiments, Sarsons et al. (2020) explore gender differences in credit for group work. They find women receive less credit for group work when employers cannot perfectly observe their contributions. Women are therefore less likely to receive tenure the more they coauthor and these differences are explained mainly by stereotypes and biases.

Erkal et al. (2020) examine whether gender contributes to biased belief formation about the leader's outcomes. They find gender distorts perceptions of outcomes in risky environments. Relative to unbiased Bayesian updating, good outcomes of women are attributed more to luck and bad outcomes of men more to effort. In treatments in which members can offer bonuses or impose penalties on decision makers (leaders), female leaders receive lower bonuses than their male counterparts. The authors identify channels underlying bonus payments and find that, while outcome matters for both men and women, beliefs about intentions only matter for men. Hence, good outcomes are necessary for women to get a bonus, but men can get a high bonus for bad outcomes as long as evaluators hold them in high regard.

4.2 Leader effectiveness

Closely related to how leaders are evaluated is the effectiveness of leaders in rallying support from their followers in order to achieve certain goals. Gender differences in leader effectiveness have been studied in two distinct ways: Firstly, by inducing leadership roles in the laboratory and secondly, by examining the performance of incumbent leaders in the corporate or political arena.[xiii]

Andreoni and Petrie (2008) find that their laboratory subjects are more willing to follow the lead of generous men in a linear Public Goods game with sequential contributions. When the first mover is male and his contribution is known with certainty, subsequent movers make more generous contributions than when a similar contribution is made by a woman. This effect increases with the proportion of men in the group.

Reuben and Timko (2018) compare the effectiveness of elected leaders and randomly selected leaders of groups playing a coordination game. Elected leaders are more effective, but teams selecting a male leader have higher performance than teams that elect a female leader, mostly because male leaders are more likely to be followed when they request the highest effort level. The difference disappears over time, as high performing leaders are more likely to be re-elected.

Grossman et al. (2015) conduct a collective action experiment with asymmetrical information in which free riding and coordination failures can prevent efficient group cooperation. Randomly selected leaders, knowing the value of the project assigned to their groups, decide first whether they want to contribute to the projects. Uninformed followers observe the leader's decision, then simultaneously decide whether to invest. There are three treatments: single-gender groups, mixed-gender groups with no leader gender signaling, and mixed-gender groups with leader gender signaling. While there are significant differences by gender in the behavior of leaders (discussed in Section 2.3), Grossman et al. report that, contingent on their leader investing, followers are as likely to invest regardless of the gender of the leader.

Comparing the effectiveness of incumbent female leaders with male leaders is often difficult to do as female leadership is rarely observed.[xiv] To circumvent this problem, researchers have used conditions created by natural policy experiments that mandate women be considered for leadership positions (for example, Norway, where publicly listed firms were mandated to have at least 40 percent woman directors, and India, where leadership positions were randomly reserved for women).[xv]

Gangadharan et al. (2016) and Gangadharan et al. (2019) are examples of how researchers can combine natural variations caused by policy (the Indian policy of reserving randomly determined village leadership roles for women) with experimental methods. Using a lab-in-the-field experiment, the two papers explore gender differences in leaders' behavior and followers' behavior toward male and female leaders. Villagers in both reserved and non-reserved villages take part in a Public Good with Leadership experiment. They find that female leaders are more likely to contribute less than what they propose, compared to men in similar positions. This negative deviation from proposals is pronounced when the leader's gender is revealed. Men contribute significantly less to the public good when the group leader is revealed to be a woman.

In summary, the studies show mixed evidence of gender differences in both the evaluation of leaders and leader effectiveness. A pattern that emerges across the papers, however, is that

perceptions of legitimacy and the social environment can be critical to facilitate the effectiveness of female leaders.

5. Interventions to address the gender gap in leadership

Despite improvements over the past decades, gender differences are evident in the workplace in terms of fewer opportunities for career advancements and lower wages for women. Decision makers in both the public and the private sectors have considered several approaches to accelerate progress toward gender parity. These include initiating diversity and leadership programs to encourage women to take up senior positions, programs to reduce unconscious biases of managers, and more explicit legislative policies, such as mandatory affirmative action.

Our focus in this section is on three broad types of interventions. First, an obvious solution to the problem of an underrepresentation of women in leadership is to simply increase female representation through mandatory affirmative action. However, such top-down measures may invite backlash, especially in the absence of wide-ranging support from peers and followers. Second, we discuss interventions aimed at addressing the willingness of women to become leaders by encouraging them to "lean in" (Sandberg, 2013). Third, we consider proposed changes to institutions that aim to debias beliefs and reduce the role of gender stereotypes for female candidates.

5.1 Interventions to regulate increases in female representation

The main purpose of affirmative action is to establish access to educational or employment opportunities for underrepresented groups, in this case, women. Several policy interventions have been initiated to close the gender gap [corporate board membership and directorships in California (McGreevy, 2018); gender-based quotas in the political arenas in Europe (Casas-Arce and Saiz, 2015); as well as the policies adopted in India and Norway (see Section 4.2)].

One of the arguments against affirmative action is that these polices prevent the allocation of jobs to the best available candidate. Testing the empirical validity of this argument using observational

data is difficult as important variables, such as merit and candidate quality, can be hard for researchers to observe and quantify.

Over the last decade several studies, using the Niederle and Vesterlund (2007) paradigm, have examined the effectiveness of affirmative action policies to enhance entry into competition. Niederle et al. (2013) explore whether affirmative action can encourage applications from female candidates who would otherwise fail to apply for positions they are actually qualified for. They find that this policy leads to a large increase in tournament entry by women and a decrease in the entry by men. The magnitude of this change is beyond what might be expected from just the changes in the probability of winning for men and women. The authors argue that the large effect on women could be explained by the reduction in the gender differences in beliefs about relative performance and by a reduction in the gender difference in the willingness to compete. They show that while some high performing men fail to enter the competition, many more high performing women do enter, hence the overall number of high performing people in the pool is unaffected.

Balafoutas and Sutter (2012) examine four policy interventions compared to a no-intervention control scenario. The policies include: 1) quotas that require a minimum fraction of winners to be women; 2) a weak preferential treatment scheme, in which a tie-breaking rule favors women in situations in which men and women have equal performance; 3) a strong preferential treatment scheme in which each woman receives a head start; and finally, 4) repeating the competition until a critical threshold of women has been reached. They find that the gender gap in the willingness to enter competition, observed in the control, is reduced with the policy interventions. The strong preferential treatment is the most effective in encouraging women to compete. The weak preferential treatment and quotas have a similar effect, and the repetition of competition generates the weakest incentives for entry. The policies influence women's behavior and do not have a significant impact on men's behavior. Importantly they also find that the policies do not lead to efficiency losses, either in terms of the selection of winners or in terms of earnings in a post-competition coordination game.

Balafoutas et al. (2016) examine if individuals are willing to support quotas and how the endogenous implementation of these policies affects performance and efficiency. An important aspect of their design is that the quota can either favor women or a completely arbitrary subgroup (defined by color: pink versus green). Subjects can vote for or against the policy, or abstain. When quotas are based on color instead of gender, performance becomes skewed; members in the advantaged (disadvantaged) color group significantly increased (decreased) their performance. The authors conclude that if there is justification for one group to receive preferential treatment, as in the case of quotas for women, but not in the case of quotas based on color, then this policy does not reduce performance or efficiency. Similarly, Ip et al. (2020) find that the effectiveness of affirmative action can depend significantly on the acceptability of the underlying reason for the initial disadvantage.

Maggian et al. (2020) ask at which stage quotas would be most effective, at the entry level or for more senior positions? They find that quotas introduced at the final stage of the competition or at both the entry and the final stages helped increase women's willingness to compete. They find that introducing a quota only at the top is as effective as introducing quotas at all stages. The interventions do not lead to any efficiency loss, as measured by performance of winning subjects and the overall earnings. However, only entry-level quotas have the perverse effect of discouraging women from competing at higher levels due to the stereotype threat (Aronson and Steele, 1995).

While affirmative action may increase the proportion of women in leadership positions, simply enforcing gender-based quotas may have little (or even a negative) effect on the acceptance of female leaders. Leibbrandt et al. (2018) address how affirmative action policies interact with other institutional features of organizations, such as a peer review process, or the social environment. They introduce two treatment dimensions, one in which an affirmative action policy is initiated and another in which subjects peer review others' performances. Introducing peer review provides an opportunity for sabotage; subjects can choose to misreport the performances of others and hence reduce others' earnings. They find that in treatments with no peer review, quotas are

effective in encouraging women to compete. In the presence of peer review, however, the quota both fails to encourage women to compete and, for those women who do compete, there is significant backlash (i.e., sabotage).

Gangadharan et al. (2016, discussed in Section 4.2) report evidence of male backlash in their lab-in the-field experiments in India. This backlash is considerably stronger in villages that were exposed to female leaders due to the affirmative action policy. Male backlash is explained by a perceived transgression of gender norms regarding greater male authority. When women become leaders, men believe their identity is violated and their behavior reflects this resentment. [xvi]

Taken together the findings from Leibbrandt et al. (2018) and Gangadharan et al. (2016) raise concerns about the effectiveness of strong policy interventions such as gender-based quotas and indicate that these may lead to perverse outcomes in some situations.

On the other hand, Beaman et al. (2009) report a positive effect of the gender quotas imposed on Indian village councils. Extended exposure to female leaders resulted in electoral gains for women. In villages required to have a female leader in the prior two elections, women were more likely to run and win elected council positions. The authors argue that this exposure weakened gender stereotypes and improved the perception of female leader effectiveness. In a follow-up paper, Beaman et al. (2012) report that this policy had the added effect of raising the career aspirations and educational attainment of adolescent girls in quota villages relative to non-quota villages.

5.2 Interventions to change women's preferences

Given the reluctance of women to be considered for leadership roles, a seemingly straightforward solution is to encourage women to "lean in." However, as discussed in Section 3.2, gender stereotypes can be especially harmful to female candidates who ask for more. Bowles et al. (2007, discussed in Section 3.2) posit that women are punished for diverging from their gender stereotype of not being greedy and not being too assertive, raising doubts about the notion that women need only to ask.

Exley et al. (2020) build on the idea that simply encouraging women to lean in to negotiations may be neglecting the fact that society is not always accepting of women who lean in. The authors take advantage of the ability to isolate mechanisms in laboratory experiments by controlling for factors that are known to contribute to the gender gap. These include the fear of backlash, ambiguity in the negotiation environment, and differences in confidence. By exogenously varying whether participants can opt out or must always enter negotiations, they compare financial gains when women choose to enter into negotiations with gains when negotiation is mandated. The results suggest that, when given a choice, women tend to avoid negotiations that would have resulted in negative profits and generally make positive profits when they do enter, thus revealing a subtle nuance in the importance of knowing *when* to lean in. A key implication of these findings is that "lean in" policies that are aimed at women may backfire in certain contexts, particularly if women face similar levels of "pushback" from their negotiating counterparts. It is therefore important to also examine the ways in which the social environment might be made more favorable to female candidates.

5.3 Interventions to change institutions

Moving away from changing women's preferences and addressing instead how institutions can be changed has also emerged as an approach to reducing gender gaps (Bohnet, 2016). Erkal et al. (2019) is a step in this direction. They start with the premise that leadership selection in most organizations often requires candidates to actively express their interest. They compare such an Opt-in mechanism with an Opt-out mechanism in which everyone qualified for the position is in the candidate pool by default, but individuals can choose to opt out of the selection process. They find that conventional Opt-in mechanisms contribute to a gender gap in participation decisions, which strikingly continues to exist even when individuals know that they are the top performers in their group. In contrast, under the Opt-out mechanism, women are more likely to choose to be a leader. They argue that by changing the selection default, an Opt-out mechanism may make being in a leadership position appear more acceptable for women. Also, participating under the Opt-out

mechanism does not necessarily convey the same image of competitiveness or aggressiveness, since one does not need to actively choose to participate. Hence, debiasing the institution, by simply changing the default from an Opt-in to an Opt-out mechanism, can be fruitful in closing the gender gap. He et al. (2019) find a similar result in the context of choosing to enter competition. In the Opt-out system, both women and men choose to compete at the same rate as men in the Opt-in system, and the Opt-out system has the additional advantage that it does not lead to penalties from evaluators making decisions about whom to hire.

Hernandez-Arenaz and Iriberri (2019b) also highlight the important role of institutions in enhancing or diminishing gender biases. In their experiment, participants complete a real-effort task and then negotiate over how the earnings should be shared in either a symmetric environment or an asymmetric environment, in which a fair split is not possible. The findings indicate that women obtain worse outcomes only in the asymmetric environment, suggesting that gender stereotypes are more likely to be activated in the absence of an institution with clear decision rules (such as a sharing rule).

The provision of direct information has been proposed as another means to the same end of debiasing beliefs. However, the evidence is not positive. Two studies ask whether direct information about risk preferences can diminish gender stereotypes. In Grossman and Lugovvsky (2011), subjects predict the lottery choice made by their fellow subjects based on visual, gender-revealing clues, responses to a risk-based survey, or both. They find that without survey information, subjects apply the gender stereotype. Survey responses alone shape guesses, but when both are provided, gender stereotypes again dominate. Grossman (2013) further explores the sequencing of information, and finds that both types of information initially affect predictions, but the gender stereotype is strong when visual cues are available, regardless of the sequence.

6. Conclusion and discussion

The experimental literature has shown that both preferences and beliefs play a role in explaining the gender gap in leadership. While differences do exist between the sexes with respect to risk

aversion, competitiveness, altruism, and willingness to lead, these differences tend to be amplified in the beliefs and perceptions about female candidates and leaders. Gender stereotypes provide crucial fuel to the flame of discrimination against women in the leader selection process. For instance, the expectation that women should not be demanding can hurt them in the process of negotiating for a promotion. Perceptions that are negatively biased by stereotypes can then feed into the evaluation of women as leaders. Being *perceived* as less effective can in fact be self-fulfilling if others are less likely to follow and work alongside female leaders. Investigating the conditions under which women are more reluctant to become leaders as well as the differences in the evaluation of male and female leaders offer interesting avenues for future work.

From a policy perspective, interventions directed at changing women's preferences or encouraging more women to lead neglect the other side of the problem, namely, beliefs that may be biased against female leaders. Without a corresponding change in general attitudes, these attempts may be ineffective at best, and at worst, women may become the targets of backlash and resentment. Greater acceptance of women in leadership roles and a greater willingness of women to participate may be aided by structural changes such as clear rules and procedures to prevent harmful stereotypes from being "activated' in the first place. These are promising areas for future research.

References

Ahern, K.R. and Dittmar, A.K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics* 127, 137-197.

Amanatullah, E.T. and Morris, M.W. 2010. Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology* 98, 256.

Andreoni, J and Vesterlund, L. 2001. Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics* 116, 293–312.

Andreoni, J. and Petrie, R. 2008. Beauty, gender and stereotypes: Evidence from laboratory experiments. *Journal of Economic Psychology* 29, pp.73-93.

Apicella, C.L., Demiral, E.E., and Mollerstrom, J. 2017. No gender difference in willingness to compete when competing against self. *American Economic Review* 107, 136-40.

Arbak, E. and Villeval, M.C. 2013. Voluntary leadership: motivation and influence. *Social Choice and Welfare* 40, 635-662.

Aronson, J. and Steele, C. M. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69, 797–811.

Babcock, L. and Laschever, S. 2009. *Women don't ask: Negotiation and the gender divide.* Princeton: Princeton University Press.

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, 714–747.

Balafoutas, L., Davis, B.J., and Sutter, M. (2016). Affirmative action or just discrimination? A study on the endogenous emergence of quotas. *Journal of Economic Behavior and Organization* 127, 87-98.

Balafoutas, L. and Sutter, M. 2012. Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science* 335, 579-582.

Baldiga, K. 2014. Gender differences in willingness to guess. *Management Science* 60, 434-448.

Banuri, S. and Keefer, P. 2016. Pro-social motivation, effort and the call to public service. *European Economic Review* 83, 139-164.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. 2009. Powerful women: does exposure reduce bias? *The Quarterly Journal of Economics* 124, 1497-1540.

Beaman, L., Duflo, E., Pande, R. and Topalova, P. 2012. Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science 335*, 582-586.

Bertrand, M., Black, S.E., Jensen, S., and Lleras-Muney, A. 2019. Breaking the glass ceiling? The effect of board quotas on female labour market outcomes in Norway. *The Review of Economic Studies* 86, 191-239.

Bertrand, M. and Duflo, E. 2017. Field experiments on discrimination. In E. Duflo and A. Banerjee (ed) *Handbook of Field Experiments* Vol. 1. Amsterdam: North-Holland. 309-393.

Beyer, S. 1990. Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology* 59, 960.

Bilén, D., Dreber, A., and Johannesson, M. 2020. Are women more generous than men? A meta-analysis. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3578038.

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

Bohnet, I. 2016. *What works: Gender equality by design.* Cambridge: Harvard University Press.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. 2007. Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics* 14, 926-937.

Booth, A. and Nolen, P. 2012a. Choosing to compete: How different are girls and boys? *Journal of Economic Behavior and Organization* 81: 542-555.

Booth, A. and Nolen, P. 2012b. Gender differences in risk behavior: Does nurture matter? *Economic Journal* 122, F56-F78.

Bordalo, P., Coffman, K., Glennaiolli, N., and Shleifer, A. 2019. Beliefs about gender. *American Economic Review* 109, 739–773.

Boring, A. 2017. Gender biases in student evaluations of teaching. *Journal of Public Economics* 145, 27-41.

Born, A., Ranehill, E., and Sandberg, A. 2019. A man's world? The impact of a male dominated environment on female leadership. University of Gothenburg, Working Paper in Economics No. 744.

Bowles, H. R., Babcock, L., and Lai, L. 2007. Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes* 103, 84–103.

Brooks, A. W., Huang, L., Kearney, S. W., and Murray, F.E. 2014. Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National Academy of Sciences*, 11, 4427-4431.

Buser, T., Niederle, M., and Oosterbeek, H. 2014. Gender, competitiveness, and career choices. *The Quarterly Journal of Economics* 129, 1409-1447.

Buser, T., Peter, N., and Wolter, S.C. 2017. Gender, competitiveness, and study choices in high school: Evidence from Switzerland. *American Economic Review* 107, 125-30.

Byrnes, J.P., Miller, D.C., and Schafer, W.D. 1999. Gender differences in risk taking: a meta-analysis. *Psychological Bulletin* 125, 367.

Carli, L.L. 2001. Gender and social influence. *Journal of Social Issues* 57, 725-741.

Carpenter, J., Frank, R., and Huet-Vaughn, E. 2018. Gender differences in interpersonal and intrapersonal competitive behavior. *Journal of Behavioral and Experimental Economics* 77, 170-176.

Casas-Arce, P. and Saiz, A. 2015. Women and power: Unpopular, unwilling, or held back? *Journal of Political Economy* 123, 641-669.

Cason, T., Gangadharan, L., and Grossman, P.J. 2020. Gender, beliefs, and coordination with externalities. Monash University working paper.

Cassar, A., Wordofa, F., and Zhang, Y.J. 2016. Competing for the benefit of offspring eliminates the gender gap in competitiveness. *Proceedings of the National Academy of Sciences* 113, 5201-5205.

Castillo, M., Petrie, R., and Torero, M. 2010. On the preferences of principals and agents. *Economic Inquiry* 48, 266-273.

Chakraborty, P. and Serra, D. 2019. Gender differences in top leadership roles: Does worker backlash matter. Working paper, Department of Economics, Southern Methodist University.

Chattopadhyay, R. and Duflo, E. 2004. Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica* 72, 1409-1443.

Chauvin, K.W. and Ash, R.A. 1994. Gender earnings differentials in total pay, base pay, and contingent pay. *ILR Review* 47, 634-649.

Chen, J. and Houser, D. 2019. When are women willing to lead? The effect of team gender composition and gendered tasks. *Leadership Quarterly* 30, 101340.

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

Coffman, K.B., Collis, M., and Kulkarni, L. 2019. *When to apply?* Working Paper.

Coffman, K.B., Exley, C.L., and Niederle, M. 2020. The role of beliefs in driving gender discrimination. Working Paper 18-054, Harvard Business School.

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

De Paola, M., Ponzo, M., and Scoppa, V. 2017. Gender differences in the propensity to apply for promotion: evidence from the Italian Scientific Qualification. *Oxford Economic Papers* 69, 986-1009.

Dittrich, M., Knabe, A., and Leipold, K. 2014. Gender differences in experimental wage negotiations. *Economic Inquiry* 52, 862–873.

Dreber, A., von Essen, E., and Ranehill, E. 2014. Gender and competition in adolescence: task matters. *Experimental Economics* 17, 154-172.

Eagly, A.H., Karau, S.J., and Makhijani, M.G. 1995. Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin* 117, 125–45.

Eckel, C. and Candelo, N. 2020. How to tame lab-in-the-field experiments. In J.N. Druckman and D.P. Green (ed.) *Advances in Experimental Political Scien*ce. Cambridge: Cambridge University Press.

Eckel, C., de Oliveira, A.C.M., and Grossman, P.J. 2008. Gender and negotiation in the small: Are women (perceived to be) more cooperative than men? *Negotiation Journal* 24, 429–445.

Eckel, C.C. and Grossman, P.J. 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior* 23, 281-295.

Eckel, C.C. and Grossman, P.J. 2008. Men, women and risk aversion: Experimental evidence. In C.R. Plott and V.L. Smith (ed.) *Handbook of Experimental Economics Results*. Amsterdam: North Holland/Elsevier Press. 1061-1073.

England, P., Budig, M., and Folbre, N. 2002. Wages of virtue: The relative pay of care work. *Social Problems* 49, 455-473.

Eriksson, K.H. and Sandberg, A. 2012. Gender differences in initiation of negotiation: Does the gender of the negotiation counterpart matter? *Negotiation Journal* 28, 407-428.

Erkal, N., Gangadharan, L., and Koh, B.H. 2020. Gender biases and performance evaluation: Do outcomes matter more than intentions. Working Paper.

Erkal, N., Gangadharan, L., and Xiao, E. 2019. Leadership selection: Can changing the default break the glass ceiling? Monash University Working Paper.

Ertac, S., Gumren, M., and Gurdal, M.Y. 2019. Demand for decision autonomy and the desire to avoid responsibility in risky environments: Experimental evidence. *Journal of Economic Psychology*, 102200.

Ertac, S. and Gurdal, M.Y. 2012. Deciding to decide: Gender, leadership and risk-taking in groups. *Journal of Economic Behavior and Organization* 83, 24-30.

Exley, C.L. and Kessler, J. 2019. The gender gap in self-promotion. *NBER Working Paper* No. w26345.

Exley, C.L., Niederle, M., and Vesterlund, L. 2020. Knowing when to ask: The cost of leaning in. *Journal of Political Economy* 128, 816–854.

Falk, A. and Hermle, J. 2018. Relationship of gender differences in preferences to economic development and gender equality. *Science 362*(6412).

Filippin, A. and Crosetto, P. 2016. A reconsideration of gender differences in risk attitudes. *Management Science* 62, 3138-3160.

Folbre, N. (ed). 2012. *For Love or Money: Care Provision in the United States*. Russell Sage Foundation.

Gangadharan, L., Jain, T., Maitra, P., and Vecci, J. 2016. Social identity and governance: The behavioral response to female leaders. *European Economic Review* 90, 302-325.

Gangadharan, L., Jain, T., Maitra, P., and Vecci, J. 2019. Female Leaders and their Response to the Social Environment. *Journal of Economic Behavior and Organization* 164, 256-272.

Gillet, J., Cartwright, E., and Van Vugt, M. 2011. Selfish or servant leadership? Evolutionary predictions on leadership personalities in coordination games. *Personality and Individual Differences* 51, 231-236.

Gneezy, U., Leonard, K.L., and List, J.A. 2009. Gender differences in competition: Evidence from a matrilineal and patriarchal society. *Econometrica* 77, 1637-1664.

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

Goldin, C. and Rouse, C., 2000. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review* 90, 715-741.

Grossman, P.J. 2013. Holding fast: The persistence and dominance of gender stereotypes. *Economic Inquiry* 51, 747-763.

Grossman, P.J., Eckel, C., Komai, M. and Zhan, W. 2019. It pays to be a man: Rewards for leaders in a coordination game. *Journal of Economic Behavior and Organization* 161, 197-215.

Grossman, P.J., Komai, M. and Jensen, J.E. 2015. Leadership and gender in groups: An experiment. *Canadian Journal of Economics/Revue canadienne d'économique* 48, 368-388.

Grossman, P.J. and Lugovskyy, O. 2011. An experimental test of the persistence of gender-based stereotypes. *Economic inquiry* 49, 598-611.

He, J., Kang, S., and Lacetera, N. 2019. Leaning in or not leaning out? Opt-out choice framing attenuates gender differences in the decision to compete. NBER Working Paper No. 26484.

Hernandez-Arenaz, I. and Iriberri, N. 2019a. A review of gender differences in negotiation. Oxford Research Encyclopedia of Economics and Finance. https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-464.

Hernandez-Arenaz, I. and Iriberri, N. 2019b. Gender differences in alternating-offer bargaining: An experimental study. Retrieved from: https://www.socialscienceregistry.org/trials/.

Holt, C.A. and Laury, S.K., 2002. Risk aversion and incentive effects. *American Economic Review* 92, 1644-1655.

Inzlicht, M. and Ben-Zeev, T. 2000. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science* 11, 365-371.

Ip, E., Leibbrandt, A, and Vecci, J. 2020. How do gender quotas affect workplace relationships? Complementary evidence from a representative survey and labor market experiments. *Management Science* 66, 805-822.

Jianakoplos, N.A. and Bernasek, A. 1998. Are women more risk averse? *Economic Inquiry* 36, 620–30.

Jiang, B. and Capra, C.M. 2018. Are (active) entrepreneurs a different breed? *Managerial and Decision Economics* 39, 613-628.

Johnson, S.K., Murphy, S.E., Zewdie, S., and Reichard, R.J. 2008. The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes* 106, 39-60.

Jung, S., Choe, C., and Oaxaca, R.L. 2018. Gender wage gaps and risky vs. secure employment: An experimental analysis. *Labour Economics* 52, 112-121.

Kamas, L. and Preston, A. 2018. Competing with confidence: The ticket to labor market success for college-educated women. *Journal of Economic Behavior and Organization* 155, 231-252.

Kanthak, K. and Woon, J. 2015. Women don't run? Election aversion and candidate entry. *American Journal of Political Science* 59, 595-612.

Leibbrandt, A., and List, J.A. 2015. Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science* 61, 2016–2024.

Leibbrandt, A., Wang, L.C., and Foo, C. 2018. Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women. *Management Science* 64, 3501-3516.

Li, S., Chaudhuri, A., and Sbai, E. 2020. (Un)Willing to lead? Men, women and the leadership gap. University of Auckland Working Paper.

Maggian, V., Montinari, N., and Nicolò, A. (2020). Do quotas help women to climb the career ladder? A laboratory experiment. *European Economic Review* 123, 103390.

McGreevy, P. September 30. 2018. Gov. Jerry Brown signs bill requiring California corporate boards to include women. *Los Angeles Times*, Retrieved from https://www.latimes.com/politics/la-pol-ca-governor-women-corporate-boards-20180930-story.html.

Moss-Racusin, C.A., Dovidio, J.F., Brescoll, V.L., Graham, M.J., and Handelsman, J. 2012. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Science* 109, 16474–16479.

Murray, R. 2008. The power of sex and incumbency: A longitudinal study of electoral performance in France. *Party Politics* 14, 539-554.

Niederle, M. 2015. Gender. In J.H. Kagel and A.E. Roth (ed.) *The Handbook of Experimental Economics* Volume 2 (481-716). Princeton: Princeton University Press.

Niederle, M. 2017. A gender agenda: A progress report on competitiveness. *American Economic Review: Papers and Proceedings* 107, 115–119.

Niederle, M., Segal, C., and Vesterlund, L. 2013. How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science* 59, 1–16.

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

O'Gorman, R., Henrich, J., and Van Vugt, M. 2008. Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society of London B: Biological Sciences* 276, 323-329.

Ortmann, A. and Tichy, L.K. 1999. Gender differences in the laboratory: evidence from prisoner's dilemma games. *Journal of Economic Behavior and Organization* 39, 327-339.

Petrie, R. and Segal, C. 2015. Gender differences in competitiveness: The role of prizes. Downloaded from: https://papers.ssrn.com/sol3/Data_Integrity_Notice.cfm?abid=2520052.

Preece, J. and Stoddard, O. 2015. Why women don't run: Experimental evidence on gender differences in political competition aversion. *Journal of Economic Behavior and Organization* 117, 296-308.

Propp, K.M. 1995. An experimental examination of biological sex as a status cue in decision-making groups and its influence on information use. *Small Group Research* 26, 451-474.

Reuben, E., Rey-Biel, P., Sapienza, P., and Zingales, L. 2012. The emergence of male leadership in competitive environments. *Journal of Economic Behavior and Organization* 83, 111-117.

Reuben, E., Sapienza, P., and Zingales, L., 2015. Taste for competition and the gender gap among young business professionals (No. w21695). National Bureau of Economic Research.

Reuben, E. and Timko, K., 2018. On the effectiveness of elected male and female leaders and team coordination. *Journal of the Economic Science Association* 4, 123-135.

Ridgeway, C.L. 2001. Gender, status, and leadership. *Journal of Social Issues* 57, 637–55.

Ridgeway, C.L., Johnson, C., and Diekema, D. 1994. External status, legitimacy, and compliance in male and female groups. *Social Forces* 72, 1051–77.

Sandberg, S. (2013). *Lean In: Women. Work and the Will to Lead*, London: WH Allen.

Sarsons, H., Gerxhani, K., Reuben, E., and Schram, A. 2020. Gender differences in recognition for group work. *Journal of Political Economy*, forthcoming.

Shurchkov, O. and Eckel, C.C. 2018. Gender differences in behavioral traits and labor market outcomes. In S.L. Averett, L.M. Argys, and S.D. Hoffman (ed.) *The Oxford Handbook of Women and the Economy*. Oxford: Oxford University Press.

Small, D.A., Gelfand, M., Babcock, L., and Gettman, H. 2007. Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation. *Journal of Personality and Social Psychology* 93, 600–613.

Taub, A. 2020. Why are women-led nations doing better with Covid-19? *The New York Times.* https://www.nytimes.com/2020/05/15/world/coronavirus-women-leaders.html.

Tonin, M. and Vlassopoulos, M. 2015. Are public sector workers different? Cross-European evidence from elderly workers and retirees. *IZA Journal of Labor Economics* 4, 11.

Endnotes

[i] Falk and Hermle (2018) report evidence suggesting that gender differences in preferences are positively related to economic development and gender equality.

[ii] And particularly relevant field experiments.

[iii] For examples of commonly used incentivized elicitation measures, see Gneezy and Potters (1997), Holt and Laury (2002), and Eckel and Grossman (2002). Studies using the Holt and Laury (2002) elicitation task typically find no gender difference (Filippin and Crosetto, 2016). Booth and Nolan (2012b) report that girls who attended same-sex schools were not more risk averse than boys. Many

studies have assessed the external validity of risk aversion measures, correlating them with a variety of real-world, high-stakes decisions as discussed in the main text.

[iv] However, their sample included only participants in entrepreneurship-related conferences and courses, and overall showed a high degree of risk tolerance.

[v] In some settings, women are not less competitive, including when competing against own past performance (Apicella et al., 2017, Carpenter et al., 2018); when the reward is high enough (Petrie and Segal, 2015); when looking at a stereotypically female task (Dreber et al., 2014); or when competing for the benefit of their children (Cassar et al., 2016). In matriarchal societies, girls are observed to be as competitive as boys (Gneezy et al., 2009). Girls who attended same-sex schools are also observed to be as competitive as boys (Booth and Nolan, 2012a).

[vi] O'Gorman et al. (2008) find that leaders with punishment powers in public goods games sacrifice their own earnings and use their power to increase group cooperation, while Gillet et al. (2011) report that leaders (i.e., first movers) in coordination games are less selfish and more altruistic, earning less than their followers.

[vii] Women have been shown to be more cooperative than men in the prisoners' dilemma game (Ortmann and Tichy, 1999).

[viii] Lower confidence is not innocuous with respect to performance. For example, women are less willing to guess on multiple choice tests, which may reflect their lower confidence level, or perhaps greater risk aversion, depending on the grading scheme (Baldiga 2014). This difference can account for much of the gender difference in scores on standardized tests.

[ix] De Paola et al (2017) document a similar pattern in the willingness of female Italian academics to put themselves up for promotion. In fields with more "objective" productivity measurement, this tendency is less likely to be observed.

[x] For a review of field experiments on this topic, see Bertrand and Duflo (2017).

[xi] Statistical discrimination is a subset of belief-based discrimination. The distinction is where the beliefs are accurate (Bertrand and Duflo, 2017).

[xii] Leibbrandt and List (2015) report similar results from a natural field experiment in which job seekers are randomly assigned to a job advertisement that either explicitly stated that wages are negotiable or do not mention that wages are negotiable.

[xiii] The early psychology literature on gender differences in leader effectiveness provides a rich context to these studies (Eagly et al., 1995; Ridgeway et al., 1994; Ridgeway, 2001).

[xiv] In recent experience, countries led by women appear to be more proactive and successful in enacting measures to combat Covid-19 (Taub, 2020). Though it is difficult to draw causal conclusions, this suggests different situations may require different leadership styles.

[xv] See, for example: Norway (Ahern and Dittmar, 2012; Bertrand et al., 2019); India (Chattopadhyay and Duflo, 2004; Beaman et al., 2009). See Section 5.1.

[xvi] Increased exposure to female leaders reduces male bias against female leaders to some extent.