Leadership and Gender Discrimination: Experimental Evidence from Ethiopia

Shibiru Ayalew   Shanthi Manian   Ketki Sheth

October 27, 2017

Abstract

Globally, women are underrepresented in top management. Using a novel "lab in the field" experiment in Ethiopia, we test whether leader gender influences the way subjects respond to leadership, and whether providing information about the leader's underlying ability changes this gender gap. We find evidence for discrimination against female leaders when subjects are given no information about leader's ability: subjects are less likely to follow female leaders. In contrast, when the leader is presented as highly trained and competent, the gender gap is reversed: subjects are *more* likely to follow women than men. The findings are consistent with a dynamic model of discrimination in which women are, on average, less likely to obtain signals of high ability. An implication of the model is that we should not observe gender discrimination among the highly educated in Ethiopia, where there are large gender gaps in educational attainment. Consistent with this, we demonstrate a lack of discrimination in a resume evaluation experiment for a management position. And, using a large sample of university administrative employees, we show that there is no gender wage gap among the highly educated.

1 Introduction

Globally, women are underrepresented in top management: for example, women hold just 17 percent of board directorships in the world's 200 largest companies (African Development Bank, 2015). In addition to equity considerations, these gaps suggest underutilization of the productivity potential of the labor force. Policy solutions have often focused on supply-side differences between male and female candidates. In high income countries, the notion that women are less likely to "lean in" and go after management positions, as well as the idea that women tend to have different leadership styles, are often discussed (Sandberg and Scovell, 2013; Eagly, 2013). Meanwhile, in low income countries, policy responses have largely tried to address lower educational attainment and skill accumulation among women (African Development Bank, 2015).

However, less evidence exists on whether individuals in the workforce respond differently to managers based solely on their gender — even when women and men are equally skilled and have similar leadership styles. If there is a differential response to female leaders, women may be less likely to succeed in management roles despite equal effort and ability. Moreover, policies that focus on altering female behavior or increasing human capital inputs to women will fall short of closing the gender gap in management.

Using a novel laboratory experiment in a sample of 304 white-collar workers in Ethiopia, we study whether individuals respond differently when they are randomly assigned to a male versus female leader. Importantly, our design allows us to hold leader ability constant: there is no direct interaction between subjects and leaders, and leader gender is the only difference between the two groups. We then cross-randomize an information treatment, where subjects are informed that the leader is highly trained and competent, to determine whether information on ability mitigates discrimination.

We find that subjects are indeed less likely to follow guidance provided by a female leader relative to a male. Surprisingly, the gender gap is not only mitigated, but is actually reversed, when subjects receive information about leader ability. Moreover, the information treatment has no effect for male leaders: the likelihood that subjects follow male leaders is statistically indistinguishable with and without information on leader ability. Our findings are inconsistent with standard models of both "taste-based" and statistical discrimination: when a male and female leader are of equal ability, as long as signals of underlying ability are unbiased, neither model can generate a reversal of the gender gap. Instead, our result suggest that the same signal must be interpreted differently for men versus women. One possible reason for this is differential barriers to entry: if subjects believe that women of equal ability are less likely to receive training, they will infer higher underlying ability when they see that a woman has been trained.

These results have real-world implications. In a labor market context, the canonical signal of ability is education. In Ethiopia, as in many places around the world, barriers to entry for women in education are well documented. For example, the World Economic Forum's 2016 Global Gender Gap Report ranked Ethiopia 132, out of 144 countries evaluated, for educational attainment. An implication of our model is that we should observe a lack of gender discrimination against women among the highly educated. In other words, when there are barriers to entry, we may *not* observe discrimination against women conditional on making it to the "top".

We move beyond the laboratory and provide evidence consistent with this implication using (i) a resume evaluation experiment and (ii) institutional data on university administration employees at an elite public university in Ethiopia.[1] In the resume experiment, we provided 264 subjects with a hypothetical job description for a top management position and a candidate resume with a randomly assigned gender listed at the top, as is customary in Ethiopia. We asked subjects to rate the resume along several dimensions (e.g., competence, likelihood of hiring, etc.). We find no differences in evaluations of candidates by gender along any dimension.

Second, we study gender wage gaps in a sample of 1,685 university administrative employees. We document an average gender wage gap of approximately 20 percent. However, we show that the wage gap is smaller at the top of the wage distribution. And we document two empirical facts consistent with our model. First, women are much less likely to have a BA or higher than men. At the same time, we find no gender wage gap among highly educated employees (those with a BA or higher).

In a large literature on gender differences in labor market outcomes,[2] this paper is one of the first, to our knowledge, to study discrimination in responses to leadership using incentivized games. Several papers have used similar methods to study gender gaps in wages and hiring, but the experimental evidence for gender discrimination is limited (Bertrand and Duflo, 2016). Psychological research, however, has documented differential responses to male and female leaders using hypothetical vignettes or trained actors (Eagly, 2013). We show that this discrimination persists when there are real stakes.

This paper also contributes to a large literature testing economic theories of discrimination[3] by documenting a set of results that cannot be explained by prevailing theories. There are two prevailing theories of discrimination in labor markets. The first, "taste-based" discrimination, is

[1]The laboratory experiment and resume evaluation experiment were conducted in subsamples of this population.
[2]Blau and Kahn (2017) review this literature
[3]See Guryan and Charles (2013) for a review.

based on a simple dislike of hiring or promoting individuals of a certain group (Becker, 1957). In contrast, "statistical" discrimination (Aigner and Cain, 1977) occurs when underlying ability is unobservable, and the underlying ability of two groups differs on average. The key insight of the statistical model is that discrimination can occur even when signals of ability are unbiased. However, neither of these theories can generate the *reversal* of the gender gap that we find in our experiment. A dynamic theory of discrimination, however, can reconcile our results: ability signals will be interpreted differently if it is more difficult for women to obtain them in the first place. Bohren, Imas and Rosenberg (2017) propose a dynamic theory of discrimination and provide evidence for a reversal of the gender gap, similar to our results, in an online field experiment.

The vast majority of the literature to date focuses on developed countries, but gender gaps in wages and management representation are a global problem. The global gender wage gap for wage and salaried workers is estimated at 23 percent (ILO, 2016). And while women hold 17 percent of board directorships globally, the analogous figures in Africa, Asia and Latin America are 14 percent, 10 percent, and 6 percent respectively. Given substantial variation in gender norms and gender disparities around the world, we may not expect the patterns of gender gaps to be the same across different settings. Indeed, our analysis of institutional data suggests important differences: while the gender wage gap is larger at the top of the wage distribution in the United States, we see the opposite in Ethiopia.

Our setting is a low-income country in which gender disparities are well documented, begin early in life, and are particularly large, as previously described. Our model suggests that this context may be important to explaining the reversal of the gender gap that we observe in our experiment, as well as the lack of gender gaps among the highly educated in the resume experiment and institutional data. One reason that signals of ability may be interpreted differently as a function of gender is that it is harder for women to obtain those signals of ability. Our results thus highlight the importance of conducting studies on discrimination in various settings, instead of assuming studies from high-income countries can be automatically applied to lower-income settings.

The rest of the paper proceeds as follows. In Section 2, we provide a theoretical framework to motivate our experiment. Section 3 provides details on the design of the three components of the study: the leadership game, resume evaluation experiment, and institutional data analysis. In Section 4, we present the results. Section 5 concludes and discusses policy implications of the results.

2 Theory

We study an employee's decision to follow the advice of either a male or a female manager. We assume that both the male and female manager have equal underlying ability θ. However, we allow mean ability in the population to vary by gender $g \in \{m, f\}$, so $\theta \sim N(\bar{\theta}_g, \sigma_\theta^2)$.[4] We further assume that both the female and the male manager are of high ability, so $\theta \geq \bar{\theta}_g$ for all g.

The employee does not observe the manager's ability. We first consider a base case in which the employee has no information about the manager except gender. Thus, the employee forms a belief $\tilde{E}(\theta|g)$ and chooses her action based on that belief. If she chooses to follow the manager's advice, she receives payoffs according to a continuous and increasing function $f(\tilde{E}(\theta|g))$. We also allow the employee's utility to depend directly on the manager's gender, as in a model of "taste-based" discrimination (Becker, 1957). Thus, the employee has utility function $u(g, f(\tilde{E}(\theta|g)))$. The employee will follow the manager if her utility is higher from doing so than from not doing so. We standardize the expected utility of not following the manager to 0, so the condition for following the manager is:

$$u(g, f(\tilde{E}(\theta|g))) > 0$$

For simplicity, we assume that utility is linear in payoffs, and that taste-based utility and utility from payoffs are additively separable. This yields $u(g, f(\tilde{E}(\theta|g))) = f(\tilde{E}(\theta|g)) - c(g)$, where c is the "taste-based" cost associated with following each gender. The employee will then follow her manager's advice if:

$$f(\tilde{E}(\theta|g)) > c(g)$$

Thus, the employee will follow the manager if the expected payoff from following the manager exceeds the taste-based cost of following the manager. From this condition, we can see our first result:

Proposition 1 *Employees are less likely to follow female managers if $c(f) > c(m)$, if $\bar{\theta}_f < \bar{\theta}_m$, or both.*

[4]Given large differences in educational attainment between men and women in Ethiopia, for example, it may make sense to assume that mean ability is higher among men.

In the absence of any other information about the manager, both taste-based discrimination and statistical discrimination result in employees being less likely to follow the female manager relative to the male manager. If there is taste-based discrimination against women, then the expected payoff from following the manager must be higher for the female manager than the male manager to compensate for the distaste. If there is statistical discrimination against women (i.e., $\bar{\theta}_f < \bar{\theta}_m$), employees are less likely to follow the female manager because the expected payoff from doing so is simply lower.

Ability signal

We now consider the possibility of introducing additional information about manager ability. Let s be a noisy but unbiased signal of ability: $s = \theta + u$, where u is independent of θ and is normally distributed with mean zero: $u \sim N(0, \sigma_u^2)$. Note that since the male and female manager are truly of equal ability, the distribution of s is the same for them both. We assume Bayesian updating and obtain:

$$\tilde{E}(\theta|s,g) = \frac{\sigma_u^2}{\sigma_s^2}\bar{\theta}_g + \frac{\sigma_\theta^2}{\sigma_s^2}s$$

In other words, when there is an additional signal of ability, employees form beliefs by taking a weighted average of the gender mean of ability and information derived from the signal. The weights depend on the relative noise of gender versus the ability signal: if gender is a noisier signal, the ability signal will be given more weight, whereas if the ability signal is noisier, the gender mean of ability will be given more weight. Now the employee will follow the manager if:

$$f\left(\frac{\sigma_u^2}{\sigma_s^2}\bar{\theta}_g + \frac{\sigma_\theta^2}{\sigma_s^2}s\right) > c(g)$$

Proposition 2 *Employees who have more information about the manager's true high ability are weakly more likely to follow the manager.*

This follows from the fact that both the male and female manager are truly of high ability, and the signal is unbiased, so $E(s) = \theta \geq \bar{\theta}_g$ for all g. Since true ability is weakly higher than the gender mean of ability for both genders, the signal will weakly increase the employee's belief about manager ability, on average. Note, however, that if $\theta = \bar{\theta}_g$, the signal will have no effect on the employee's behavior, since the signal is providing no new information.

Proposition 3 *An unbiased signal of high ability can reduce the gender gap in beliefs about ability, but cannot reverse it.*

The gender gap in beliefs about ability is $\frac{\sigma_u^2}{\sigma_s^2}\bar{\theta}_m + \frac{\sigma_\theta^2}{\sigma_s^2}s - \frac{\sigma_u^2}{\sigma_s^2}\bar{\theta}_f - \frac{\sigma_\theta^2}{\sigma_s^2}s$. This reduces to $\frac{\sigma_u^2}{\sigma_s^2}(\bar{\theta}_m - \bar{\theta}_f)$. The presence of any signal reduces the gender gap in beliefs, since $\frac{\sigma_u^2}{\sigma_s^2}$ is necessarily less than 1. The gender gap is further reduced as the precision of s increases and σ_u^2 decreases. If s is a perfect signal of ability ($s = \theta$), the gender gap is eliminated. However, since $\frac{\sigma_u^2}{\sigma_s^2} > 0$, an unbiased signal can generate a reversal in beliefs (i.e., a negative gender gap) only if $\bar{\theta}_m - \bar{\theta}_f < 0$, which contradictions our initial assumption.

Signal differing by gender

We now consider the scenario where the employee's interpretation of the signal is *not* unbiased and maps differently to each gender. We show that this can generate of reversal of beliefs about ability even when $\bar{\theta}_f < \bar{\theta}_m$. This result is consistent with a dynamic model of discrimination (Bohren, Imas and Rosenberg, 2017).

Let $s = \theta - \gamma_g + u$, for some constant γ_g, where $\gamma_m = 0$ and $\gamma_f > 0$. Therefore, for the same level of ability, the employee assumes that women will produce, on average, a lower signal than men. This is consistent with the idea that is more difficult for women to obtain education or training.

Proposition 4 *If the signal mean differs by gender, then it is possible for the signal s to reverse the baseline gender gap in beliefs about ability.*

Under the new specification of s, we have:

$$\begin{aligned}
E(\theta|s,g) &= E(\theta) + \frac{\sigma_\theta^2}{\sigma_s^2}[s - E(s)] \\
&= \bar{\theta}_g + \frac{\sigma_\theta^2}{\sigma_s^2}[s - \bar{\theta}_g] + \gamma_g] \\
&= \frac{\sigma_u^2}{\sigma_s^2}\bar{\theta}_g + \frac{\sigma_\theta^2}{\sigma_s^2}[s + \gamma_g]
\end{aligned}$$

The gender gap is now $\frac{\sigma_u^2}{\sigma_s^2}(\bar{\theta}_m - \bar{\theta}_f) + \frac{\sigma_\theta^2}{\sigma_s^2}(\gamma_m - \gamma_f)$. This can be negative if $\frac{\sigma_u^2}{\sigma_s^2}(\bar{\theta}_m - \bar{\theta}_f) <$

$\frac{\sigma_\theta^2}{\sigma_s^2}(\gamma_f - \gamma_m)$. Thus, individuals viewing the same signal for men and women will conclude that it indicates higher ability for the female leader, on average. If the barriers to women obtaining the signal are severe enough, this effect can reverse the gender gap in beliefs.

3 Study Design

We conducted the study in Adama, Ethiopia, in a sample of full-time administrative employees at Adama Science and Technology University (ASTU). Our primary results are based on an experiment we conducted in a subsample of these employees. We supplement the experimental results with data from a survey we conducted as well as institutional human resources data on the universe of ASTU administrative employees.

3.1 Context

Ethiopia generally performs poorly on global indicators of gender inequality. For example, in the World Economic Forum's 2016 Global Gender Gap Report, Ethiopia ranked 109 of 144. This low rank was driven by their rank on sub-indexes related to education and labor market outcomes: they ranked 106 on "Economic participation and opportunity" and 132 on educational attainment. However, the country has instituted a number of affirmative action policies designed to reduce gender gaps. In 2016, as part of its annual Country Policy and Institutional Assessment (CPIA) exercise, the World Bank assigned Ethiopia a Gender Equality Rating of 3 on a scale of 1 (low) to 6.[5]

Adama Science and Technology University (ASTU) is an elite public university located about 100 km from the capital, Addis Ababa. Table 1 shows summary statistics for all administrative employees at ASTU, based on institutional data from the human resources department. Educational attainment in the sample is high: on average, employees completed 12 years of education, which corresponds to secondary school completion. In contrast, in the Ethiopian population more broadly, 48.3 percent females and 45.7 percent males are out of secondary school (World Bank, 2017). Nearly 30 percent of the sample has a BA or higher, while the gross tertiary enrollment ratio in Ethiopia

[5]The gender equality ranking assesses the extent to which the country has installed institutions and programs to enforce laws and policies that promote equal access for men and women in education, health, the economy, and protection under law.

Table 1: Summary Statistics

	(1) Total	(2) Male	(3) Female	(4) Diff.
Female	0.56			
	(0.50)			
Tenure	8.00	7.61	8.31	-0.71*
	(5.55)	(5.95)	(5.20)	
Years of education	12.87	13.04	12.73	0.31*
	(3.01)	(3.23)	(2.83)	
BA or higher	0.30	0.38	0.23	0.14***
	(0.46)	(0.48)	(0.42)	
MA or higher	0.02	0.04	0.01	0.03***
	(0.15)	(0.20)	(0.09)	
Salary	2354.62	2629.83	2135.97	493.85***
	(1536.24)	(1878.60)	(1151.46)	
Observations	1685	746	939	1685

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard deviations in parentheses.

is just 8 percent (World Bank, 2017). Turnover among administrative employees at ASTU is low: average job tenure is 8 years.

Women represent 56 percent of the sample, which suggests that they are over-represented in the sample, but only slightly. In 2012, women and men with an advanced education were almost equally likely to be in the labor force (although the female labor force participation rate is about 15 percentage points lower overall). We observe significant differences in job tenure by gender: women have been with the institution longer.

Importantly for the interpretation of our model, women in the sample have significantly fewer years of education and are 37 percent less likely to hold a BA. They are also 75 percent less likely to hold a Masters degree. Though we were unable to find comparable national statistics on education, this does mirror the general trend of gender gaps in education completion in Ethiopia. For example, in primary and secondary school, the gender parity index of school enrollment is 1. But for tertiary school, the gross enrollment gender parity index is .5 (World Bank, 2017).

3.2 Leadership Game

3.2.1 Sample

For the leadership game, we selected a subsample of the university administrative employees that hold a BA or higher.[6] Using a list of employees provided by the human resource department, we contacted all administrative employees with a BA or higher ($n = 500$), and implemented the game until we reached 150 female subjects and 150 male subjects. The game was ultimately implemented over 6 days. Most eligible subjects who did not participate (about 40 percent) were excluded because we were unable to locate them during the week of the study.

3.2.2 Overview of design

The leadership game was designed to test two questions, as previously described: whether individuals respond differentially to male and female leaders *holding all else constant*; and whether information about the leader's ability changes the gender gap. The basic setup of the game is that subjects are randomly assigned to either a male or female "leader", and the subject-leader pair completes several tasks. However, there is no direct interaction between the subject and the leader, allowing us to hold the leader's behavior constant across male and female leaders.[7] The subject is given some information about their leader: their leader's gender, as well as their leader's age range, and that their leader works in a similar position at a different university. In general, we are interested in the likelihood of subjects following the guidance provided by their leaders as a function of their leader's gender, and whether any gender gap can be mitigated by providing information about the leader's ability.

The game consists of two tasks: a logic game (Tower of Hanoi) and a signaling game adapted from Cooper and Kagel (2005). The primary purpose of the first task is to elicit beliefs on performance, and the primary purpose of the second task is to measure whether subjects follow their leader's directions.

In the logic game, subjects are shown a Tower of Hanoi and are asked to move the tower from

[6] We restricted the game to highly educated employees because we wanted to focus on white collar workers, and because we believed that the game may be too complicated for subjects with low levels of education.

[7] The leaders were real individuals at another University who actually played the games as described to the subjects. To hold behavior constant, the leaders played ahead of time, and we only matched subjects to a subset of male and female leaders who played in the same way and had the same outcomes.

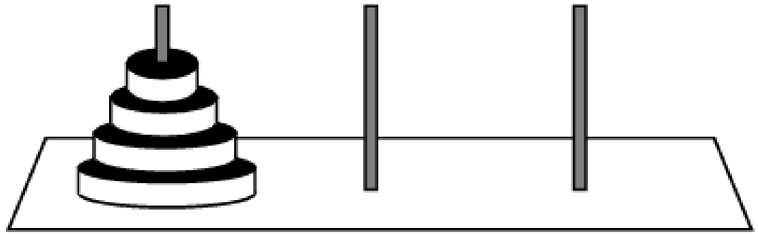

Figure 1: Tower of Hanoi

one pole to another (see Figure 1). They can only move one disk at a time, and a larger disk cannot be placed on a smaller disk. Subjects are shown the game and allowed to practice once with three disks. The subject is then told that they will be asked to solve the Tower using four disks and that the minimum moves are 15.

Prior to actually playing, we asked subjects how many moves they think *they* will require to move the tower, how many moves they think *the leader* will require to move the tower, and finally how many moves they think the leader guessed *they* would require to move the tower.[8] The subject is then given one chance to move the tower in a time limit of four minutes. The subject was paid as a decreasing function of number of moves required to move the tower and number of moves further from their guess (see Appendix Figure A for compensation schedule).

The second task was a signaling game adapted from Cooper and Kagel (2005). Their game is an example of a "Eureka" problem: there is a clear correct answer that is difficult to guess, but obvious once known. The game is a two player game in which Player 1 selects a number, and Player 2 then responds "In" or "Out" after seeing what Player 1 has selected. Player 1 is one of two types, Type A and Type B, with 50 percent probability. However, Player 1's type is unknown to Player 2. Each players' payout is determined by the choices each player has made and Player 1's type. See Figure 2 for the payout structure of the game.[9] Player 1 is always better off if Player 2 selects Out. Player 2 is better off selecting In when playing with Type A, and selecting Out when playing with Type B. If Player 2 has no information except the distribution of types in the population, it is better to play In.[10]

A naive Type B Player 1 will select 3, observing that conditional on Player 2's selection, 3 always

[8]The first two questions were randomly ordered for each subject

[9]The original game by Cooper and Kagel had 7 numbers for Player 1 to select, we adapted the game to exclude the extreme options and have only 5 numbers

[10]That is, conditional on a 50% probability of type A v.type B, the expected value of In is 350 while the expected value of Out is 250

Player 1:

Type A				Type B		
A	In	Out		B	In	Out
1	168	444		1	276	568
2	150	426		2	330	606
3	132	408		3	352	628
4	56	182		4	334	610
5	-188	-38		5	316	592

Player 2:

	Type A	Type B
In	500	200
Out	250	250

Figure 2: Signaling Game Payoffs

provides the highest payout. But such a strategy does not take into consideration that Player 1 can signal to Player 2 that he is Type B by playing Out. If Player 2 knows that Player 1 is Type B, Player 2 will play Out instead of IN. Therefore, by playing 5, Player 1 can induce Player 2 to select Out, and receive a higher expected payoff.

We were interested in only the Eureka component of the game and the strategic interaction between leaders and subjects. The research question focuses on whether leader gender and ability change the likelihood that the subject will follow the advice to play the optimal strategy of 5. We therefore assigned all subjects to be Player 1 Type B, and Player 2 was played by a computer. We programmed the computer to mimic how university students had played in Cooper and Kagel (2005). To make this clear to the subjects, they were told that the computer did not know whether they were Type A or Type B. In addition, we included the following statement: "Though you are playing a computer, the computer has been programmed to mimic how real life university students have played this game, and so the computer does not always respond in the same way to a given number."

After being introduced to the directions of the game, the subject was then asked to complete a "practice round" in which they selected which number they believe they would play, but the computer did not respond to this selection. Subjects were then asked what they believed was the probability of receiving each possible payoff in the first round. We expected non-zero probabilities on only two of the options (as the subject selects which number they will play), but the majority of subjects did include positive probabilities on more than two possible payouts. The subject was then asked what they believed was the probability of the leader receiving each possible payoff in the first round. Using these two questions, we were able to calculated the subject's belief of the expected point value for him/herself and their leader.

The subject then played 10 rounds on the game. Prior to each round, the subject is able to observe how their assigned leader played for that given round.[11] In addition, subjects are told that the leader can send them messages. [12] The messages were displayed on an android app by the enumerator (see Figure 3). The enumerator additionally recorded the leader's play and outcome for each round on a piece of paper in front of the subject.

Figure 4 provides an overview of the experiment. We completed the game in a span of 6 days. Due to subjects discussing the game with colleagues, we relabeled the choices for Day 5 and Day 6. Specifically, Player 1 selected from two different sets of letters for Days 5 and 6, and the computer responded with "left/right" and "up/down."

3.2.3 Experimental Treatments

We implemented a cross cutting randomization of two treatments: leader gender and information on leader ability. Subjects were randomly assigned into one of four groups: Female leader with no information on ability, male leader with no information on ability, female leader with information on ability, and male leader with no information on ability.[13]

[11]Leaders were selected at a different university a week prior. Unlike the subjects in the primary study, the leaders were given extensive training on how to play each task. We selected the two top performing leaders, one male and one female, to be assigned to subjects. Both of these leaders selected 5 for each round, and the Computer responded "Out" for every round. Leaders received a bonus based on the average performance of the team members assigned to them. Subjects were told that their leader's compensation is partly based on how well the subject performs on the task.

[12]To control the content of the messages, messages were pre-written and leaders selected whether or not to send the messages to the subjects. All leaders chose to send the messages.

[13]We randomized leader gender and then independently randomized the ability treatment, so the subjects are not perfectly evenly distributed across treatments. The distribution is as follows. Female leader with no information on ability: n = 78. Male leader with no information on ability: n = 71. Female leader with information on ability: n = 70. Male leader with no information on ability: n = 85.

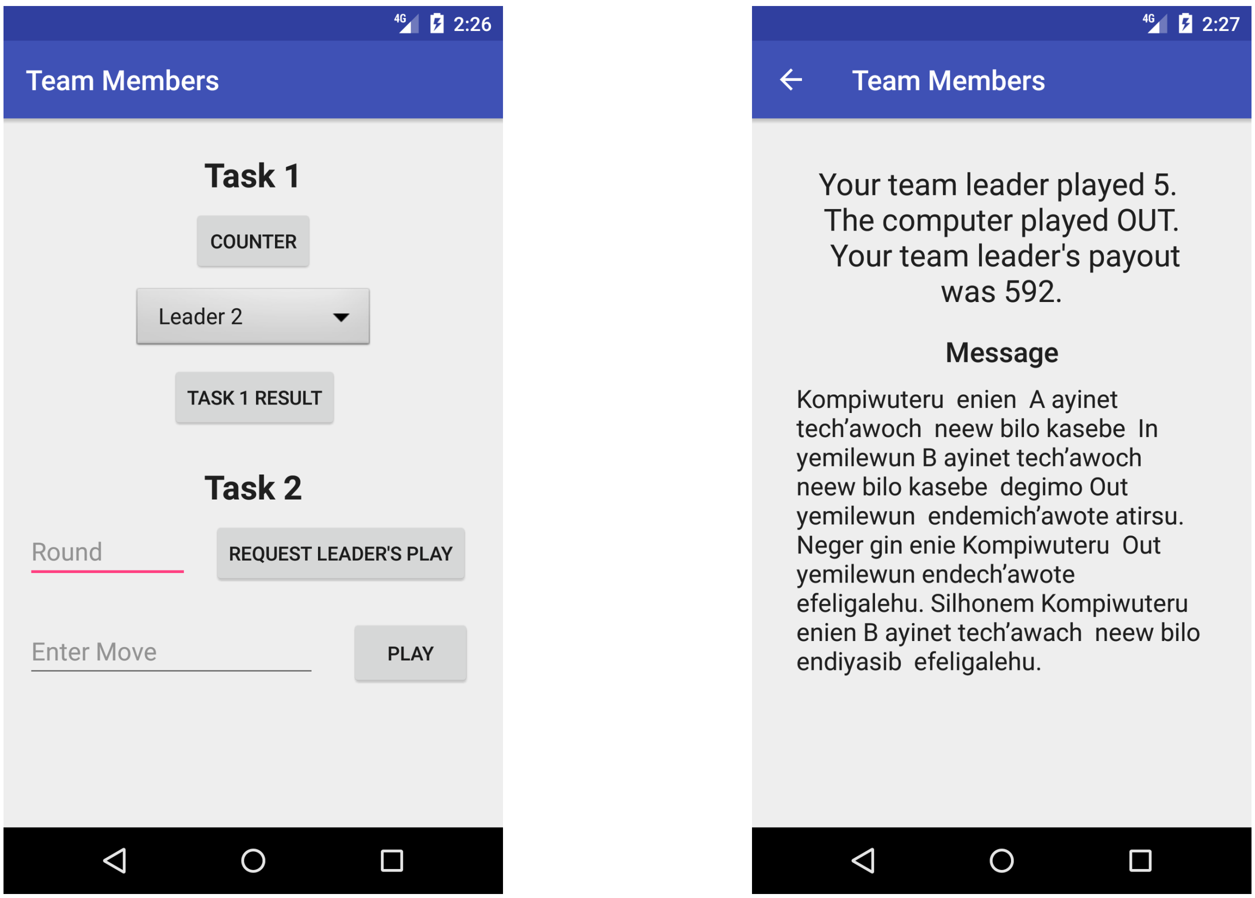

Figure 3: Leader result and messages as shown to subjects

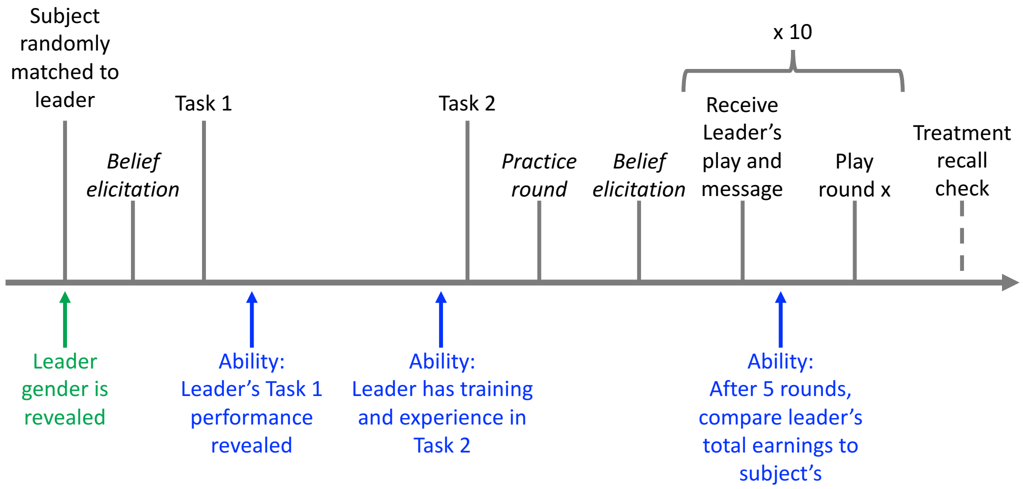

Figure 4: Timeline of Leadership Game

Leader Gender

Subjects were randomly assigned either the male leader or the female leader. Recall, the information provided to the subjects about how the leaders played are identical, and subjects do not personally interact with their leaders. This ensures that the leaders were identical to each other, except for gender. In addition to telling the subjects the gender of their leader, we provided gendered pseudonyms for the leader (mentioned 23 times in the enumerator's script) and relied on the gendered language of Amharic to make the leader's gender salient. To confirm that subjects were aware of their leader's gender, we asked subjects a series of questions on the characteristics of their leader, including gender, on the last two days of the experiment – 95 percent recalled the correct gender of the leader.

Leader Ability

Subjects were randomly assigned whether or not they were given information on the leader's ability. The ability treatment consists of three components. First, after the "Tower of Hanoi" logic game, the enumerator informed the subject that the leader completed the task in the minimum number of moves, and noted how many moves fewer this was than their own performance. [14] Second, inn the introduction to the second task, subjects were explicitly told that unlike themselves, the leader has already played the game and is an experienced player. Finally, after 5 rounds of play, the enumerator totalled the points earned by the leader versus the subject to highlight the (expected) point advantage by the leader.

3.2.4 Validity of randomization

Subjects were assigned a treatment once they arrived for the experiment. The randomization was stratified by subject gender. We had generated a random ordering of 150 treatment assignments per male and female subjects to be assigned as subjects arrived. For the last two days of the experiment, we rerandomized using a blocked randomization in groups of four, because we were concerned that we may not meet our recruitment targets (although we were ultimately successful in meeting the target). In all analyses, we account for differing randomization probabilities use inverse probability weights.

Table 2 confirms the validity of our randomization. Using information on the subjects provided by the human resources department, we confirm that subject characteristics are balanced across the four treatment groups using a linear regression of treatment assignment on each characteristic. We also confirm pairwise balance in the bottom three rows of Table 2. Column 1 confirms balance on subject, which is satisfied by construction because we stratified on subject gender. We also confirm balance on salary, job level, education, and tenure, none of which were used for stratification.

In additional to balance across subject characteristics, we may be concerned that the pseudonyms we used to connote gender also connoted other characteristics (e.g., ethnicity, age). In Ethiopia, there are significant differences in ethnicity (Amhara and Oromic are the two dominant ethnicities) and religion (Orthodox Christianity and Islam are dominant). To the extent that names connote information on ethnicity and religion, we want to confirm that our treatments are balanced across other information contained in the pseudonyms. The pseudonyms assigned to leaders were selected from a listing exercise conducted for another study in an Amharic region of Ethiopia (Ahmed and Mcintosh, 2017). We therefore oversample Oromic names in our selection. The listing exercise had

[14] Note that subjects were not informed of the extra practice and training that leaders received for the logic game, regardless of treatment assignment

Table 2: Randomization balance

	(1) Fem. subject	(2) ln(Salary)	(3) Level	(4) Years Ed.	(5) MA or higher	(6) Job tenure
Female leader only (F)	0.0173	-0.0213	-0.145	0.00175	0.00848	238.2
	(0.0817)	(0.0634)	(0.446)	(0.0813)	(0.0401)	(328.3)
Ability signal only (A)	-0.0189	-0.00813	0.151	0.0556	0.0354	71.63
	(0.0803)	(0.0597)	(0.424)	(0.0865)	(0.0427)	(335.7)
Female leader & Ability (FA)	-0.0383	-0.00636	-0.149	0.117	0.0587	-276.9
	(0.0840)	(0.0610)	(0.420)	(0.100)	(0.0494)	(342.2)
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	304	304	304	304	304	304
p-val: F = A	0.649	0.839	0.510	0.535	0.535	0.586
p-val: A = FA	0.812	0.977	0.481	0.554	0.650	0.268
p-val: F = FA	0.503	0.821	0.994	0.251	0.312	0.0959

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

also collected information on the following basic demographic information on characteristics of the person with the given name: ethnicity, religion, age, and grade completed. Table 3 confirms that the characteristics associated with the pseudonym assigned to each subject in a given treatment are balanced across treatment arms[15].

3.2.5 Estimating Equations

Our primary research question is whether women in management or leadership positions are less likely to be successful due to differential responses to their advice based solely on gender. In the leadership game, this correspond to the hypothesis that subjects are less likely to play strategically (defined as playing 4 or 5, following (Cooper and Kagel, 2005). We additionally hypothesized that information indicating the leader is trained and competent mitigates such gender gaps.

To test these hypotheses we estimate the following equation using a linear regression model:

$$R_{ir} = \alpha + \beta_1 * FL_i + \beta_2 * Ability_i + \beta_3 FL * Ability_i + \epsilon_{ir} \tag{1}$$

[15]The results are robust to the exclusion of day fixed effects.

Table 3: Pseudonym balance

	(1) Amhara	(2) Oromo	(3) Age	(4) Grade	(5) Orthodox
Female leader only (F)	-0.0188	-0.00914	0.670	0.219	-0.0220
	(0.0554)	(0.0708)	(2.365)	(0.263)	(0.0700)
Ability signal only (A)	-0.0537	-0.0104	-0.932	0.145	-0.0689
	(0.0568)	(0.0697)	(2.278)	(0.227)	(0.0665)
Female leader & Ability (FA)	-0.0265	0.00721	-0.409	0.160	-0.0477
	(0.0597)	(0.0754)	(2.517)	(0.270)	(0.0712)
Day FE	Yes	Yes	Yes	Yes	Yes
Observations	304	304	304	304	304
p-val: F = A	0.544	0.985	0.444	0.781	0.466
p-val: A = FA	0.658	0.807	0.816	0.956	0.743
p-val: F = FA	0.900	0.826	0.648	0.848	0.700

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

where R is one of two outcomes. Our primary outcome of interest is an indicator for playing strategically (i.e., selecting 4 or 5) for subject i in round r. We also study simply playing 5, since subjects may simply mimic the leader without understanding the logic behind their play. FL is an indicator for being randomly assigned a female leader, *Ability* is an indicator for being randomly assigned receipt of information on the leader's ability, and $FL * Ability$ is the interaction of the two indicators. [16] We additionally include an indicator of whether the practice round selection was equivalent to the outcome of interest, day fixed effects, and round fixed effects to increase precision of our estimates and to directly control for changes we made on the latter days of the experiment.[17] Standard errors are clustered at the individual level, corresponding to the level of randomization.

Based on our model, we hypothesized the following:

$\beta_1 < 0$: In the absence of information, directions provided by female leaders are less likely to be followed relative to directions provided by male leaders.

$\beta_2 > 0$: Providing information on the leader's ability increases the likelihood of subject's following the leader's directions.

[16] As previously described, we corrected for varying randomization probabilities using inverse probability weights. The exclusion of these weights does not qualitatively change the results.

[17] However, our results are robust to excluding these fixed effects (with a slight reduction in statistical power) and using a probit model.

$\beta_3 > 0$: Providing information on the leader's ability increases the likelihood of a subject's following the female leader's directions more than male leaders.

We additionally hypothesized the standard prediction of information in discrimination for unbiased signals of ability:

$\beta_1 + \beta_3 <= 0$: Though information mitigates gender differentials, it will fall short of eliminating the gap completely. This sum represents the gender gap conditional on ability information (i.e., the probability of following the directions of a female leader with ability information relative to a male leader with ability information.) The reversal of the gender gap predicted by a model in which signals are interpreted differently for men and women corresponds to $\beta_1 + \beta_3 > 0$.

3.3 Resume Experiment and Survey

Upon completion of the experimental game, we implemented a detailed survey that began the following week. We successfully followed up with 86.8 percent of the experimental subjects [18] The survey focused on gender in the workplace and included questions on basic demographics, household responsibilities, and attitudes about gender in the workplace.

However, prior to providing the survey questionnaire, we asked subjects to evaluate two hypothetical resumes in which the gender was randomly determined (resulting in four different resumes: female/type 1, male/type 1, female/type 2, male/type 2). Therefore, each subject received a resume for a female candidate and for a male candidate. Upon reviewing each resume, subjects evaluated the potential candidate on a increasing scale of 1 to 5 on competence, likeability, and willingness to hire. They additionally suggested a salary to be offered to the candidate. Upon reviewing both resumes, subjects were then asked to rank the candidates in terms of competence, likeability, and willingness to hire.[19]

[18] Attrition was not due to lack of consent or desire to participate, but rather driven by the difficulty in finding the same subjects by the enumerators. Because we implemented the survey over the summer, many employees were on leave.

[19] The exact questions were as follows: 1.“I will first ask you about the competency of the candidate. By competency, I mean for you to evaluate the candidate based on how well you think he will perform on the requirements of the job. Based on the resume, is his competency: poor, fair, good, very good, or excellent?“ 2. “I will now ask you about the likeability of the candidate. By likeability, I mean for you to evaluate the candidate based on how well you think he will get along with his colleagues, including the employees he will directly supervise. Based on the resume, is his likeability: poor, fair, good, very good, or excellent?” 3. “I will now ask you about how willing you would be to hire the candidate for the position. Based on the resume, would you be very unwilling, slightly unwilling, neither unwilling or willing, slightly willing, or very willing to hire him?“ 4. “If this job candidate were hired, what monthly salary would you offer him, in Ethiopian birr?”

Table 4: Resume Experiment Balance

	(1) Fem. subject	(2) Years of education	(3) ln(Salary)	(4) tenure_{days}	(5) Level
Female Resume	-0.0174	0.0620	-0.0400	401.7	-0.245
	(0.0618)	(0.0722)	(0.0454)	(246.5)	(0.324)
Resume Version	-0.0536	-0.0601	0.0219	-221.3	0.0767
	(0.0618)	(0.0722)	(0.0453)	(246.4)	(0.324)
Constant	0.528***	16.12***	8.078***	2994.0***	13.41***
	(0.0541)	(0.0631)	(0.0397)	(215.5)	(0.283)
Observations	264	264	264	264	264

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Because the survey clearly indicated the study was about gender, and we knew from our earlier experimental game subjects would discuss the content of the survey, we expected that desirability bias may be a concern for our resume evaluation experiment. We therefore controlled the ordering of how resumes were given to subjects. Each subject was given an initial resume to evaluate first (which was randomly determined to be either the male or female resume) and were not told that a second resume would be followed. Therefore, in comparing the evaluations on this first resume, we are not asking subjects to compare directly across gender, but instead relying on differences across subjects. In evaluation of the second resume, subjects may then become aware that the two resumes differ on gender and may anchor their responses to their first evaluation as a function of social desirability on gender. We do indeed find evidence of this (the second resume is evaluated more favorably if the second candidate is female and less favorably if the second candidate is male, but we find no similar differences across subjects). As a result, we exclude analysis of the second resume and focus only on the first resume where we can be more confident that social desirability bias is less likely to bias our results.

Because of uncertainty in scheduling survey interviews with subjects, we again randomized the treatment (which of the four resumes) by creating a random ordering in groups of four for each enumerator and then had them go in the order of their list when interviewing subjects. Table 4 confirms the validity of our randomization by documenting that subject characteristics were balanced across treatment arms.

3.3.1 Estimating Equation

Unlike the experimental game which focuses on whether subjects follow directions deferentially based on leader gender, the resume evaluation is a direct test for discrimination against women in the hiring process and whether the same candidate would be evaluated differently based solely on gender. We test for this using the following linear regression model:

$$Outcome_i = \alpha + \gamma_1 * FC_i + \gamma_2 * ResumeType_i + \epsilon_i \tag{2}$$

where *Outcome* is competence, likeability, hireability, or salary offer (in logs); *FC* is an indicator of whether the resume was randomly assigned to be a female candidate, *ResumeType* is a control for which resume was given; and i represents subject.

4 Results

4.1 Leadership Game

Table 5, Columns 1 and 3 show our primary results from estimating equation (2). Column 1 uses the outcome variable of strategic play (i.e., selecting 4 or 5), and Column 3 uses the outcome variable of playing 5. In both columns, we find that in the absence of ability, subjects with female leaders (and no information on ability) were 6 percentage points less likely to play in accordance with their leader's directions (see β_1). Relative to subjects with male leaders with no information on ability, this reflects a 10 percent reduction in adherence to the leader's recommendation.

Surprisingly, we find that information on ability had no effect for subjects with male leaders: subjects were equally likely to follow male leaders whether or not they were given information on the leader's experience or training (see β_2). This would suggest that for men, the signal did not provide information on ability that was greater than the expected group mean of men. In other words, the signal we provided of being capable of performing well on the tasks was already in line with the expectation of how average males would perform.

However, the information on ability does have a large effect for subjects assigned to female leaders (see β_3). Surprisingly, $\beta_1 + \beta_3 > 0$, which means that upon having information on ability, subjects were more likely to follow the directions provided by female leaders relative to male leaders.

Table 5: Leadership Game Results

Dependent Variable:	Strategic Play		Played 5	
	(1)	(2)	(3)	(4)
Fem. leader=1	-0.0590*	-0.00172	-0.0640	-0.0143
	(0.0352)	(0.0245)	(0.0407)	(0.0289)
Ability=1	-0.00301	0.0536**	0.000942	0.0501*
	(0.0350)	(0.0244)	(0.0409)	(0.0278)
Fem. leader=1 × Ability=1	0.115**		0.0998*	
	(0.0479)		(0.0562)	
Constant	0.449***	0.425***	0.123*	0.101
	(0.0559)	(0.0549)	(0.0679)	(0.0656)
Day FE	X	X	X	X
Round FE	X	X	X	X
Practice round	X	X	X	X
Observations	3020	3020	3020	3020
Control group mean		0.618		0.374
P-val.: F + F × A = 0	0.0891		0.366	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, clustered at subject level. Strategic play is defined as playing 4 or 5. 5 is the highest expected value play, and the leader played 5 in every round.

This suggests that the signal resulted in subjects inferring a higher level of ability for women than for men. Since the signal was identical for male and female leaders, and the true ability of male and female leaders was also identical by construction, the results imply that interpretation of the signal was not unbiased.

It is also worth noting that because of the reversal in discrimination caused by the ability signal, the response to female and male leaders is the same *on average*. This is shown in Columns 2 and 4: when we do not control for the interaction term, the effect of having a female leader is small and insignificant. This average effect masks the discrimination against women in the absence of information and the increased likelihood of following female leaders when ability is signaled. The finding highlights how, because of the dynamics of discrimination, estimates of discrimination may be biased towards zero and hide important heterogeneity.

Table 6 estimates our results separately for male and female subjects. Though less precise, the estimates do suggest that the reversal of discrimination may be more driven by female subjects. If women have a greater understanding of the barriers females face to attain "signals of ability", then

Table 6: Heterogeneous Treatment Effects by Subject Gender

Dependent Variable:	Strategic Play		Played 5	
	(1) M. subjects	(2) F. subjects	(3) M. subjects	(4) F. subjects
Fem. leader=1	-0.0683	-0.0600	-0.0440	-0.0927*
	(0.0488)	(0.0530)	(0.0605)	(0.0554)
Ability=1	0.0107	-0.0144	0.0828	-0.0804
	(0.0517)	(0.0481)	(0.0611)	(0.0520)
Fem. leader=1 × Ability=1	0.0979	0.135**	-0.00673	0.213***
	(0.0682)	(0.0683)	(0.0771)	(0.0799)
Constant	0.445***	0.457***	0.0904	0.173**
	(0.0665)	(0.0951)	(0.0995)	(0.0780)
Day FE	X	X	X	X
Round FE	X	X	X	X
Practice round	X	X	X	X
Observations	1560	1460	1560	1460

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, clustered at subject level. Strategic play is defined as playing 4 or 5. 5 is the highest expected value play, and the leader played 5 in every round.

it is likely that females would be more likely to infer higher levels of ability for a given signal.

Our estimates of belief expectation on how well the leader will perform in Task 2 can act as a robustness check for our results. Unfortunately, the belief expectation exercises were difficult for subjects to understand and thus were likely very noisy estimates of belief. However, as Table 7 shows, the pattern of the magnitudes of the beliefs elicited for Task 2 directly align with the pattern of being likely to follow the leader's directions in Table 5. Female leaders (relative to male leaders) were assessed to perform more poorly (i.e., lower expected value) when no information was provided on ability, but were assessed to perform more effectively (relative to male leaders) when information on ability was provided. Our results lack statistical precision and thus cannot be differentiated from having no effect on expected value of performance, but the fact that they exhibit the same pattern as our primary results are suggestive of the robustness behind our results in Table 5.

4.2 Resume Experiment and Observational Data

This section presents results from the resume evaluation experiment. As shown in Table 8, we surprisingly find no differential evaluation of resumes as a function of candidate gender. We do find

Table 7: Beliefs about leaders

Dependent Variable:	Leader's performance	
	(1)	(2)
	Task 2	Task 2
Fem. leader=1	1.304	-5.812
	(6.416)	(9.056)
Ability=1	13.50**	6.362
	(6.466)	(9.527)
Fem. leader=1 × Ability=1		14.39
		(12.98)
Constant	456.2***	459.4***
	(12.78)	(13.00)
Day FE	X	X
Observations	301	301

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 8: Resume Evaluation Results

	(1)	(2)	(3)	(4)
	Competence	Likeability	Likelihood of Hire	Log Salary
Female Resume	-0.000946	0.0392	-0.0870	-0.0400
	(0.127)	(0.113)	(0.155)	(0.0454)
Resume Type	0.246*	0.0336	-0.103	0.0219
	(0.127)	(0.113)	(0.155)	(0.0453)
Observations	263	263	263	264

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Table 9: Gender Wage Gap at Adama University

	(1) ln(Salary)	(2) ln(Salary)	(3) ln(Salary)
Female	-0.198***	-0.129***	-0.0861***
	(0.0234)	(0.0161)	(0.0197)
Tenure		0.0281***	0.0268***
		(0.00140)	(0.00168)
Years of education		0.0509***	0.0363***
		(0.00332)	(0.00402)
BA or higher		0.383***	0.337***
		(0.0262)	(0.0255)
MA or higher		0.395***	0.419***
		(0.0504)	(0.0647)
Constant	7.744***	6.701***	6.938***
	(0.0173)	(0.0403)	(0.281)
Work Unit FE	No	No	Yes
Observations	1685	1665	1665

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

that subjects are more likely to favor one type of resume, in particular with respect to competency, suggesting that subjects are paying attention to the quality of the resume when considering their responses. We additionally check for "resume comprehension" prior to asking the subject to evaluate the resume, confirming that candidates were aware of the gender of the candidate (95 percent correctly identified the candidate's resume). Therefore, our lack of discrimination results suggests that though subjects were aware of the candidate's gender and were thoughtful about their responses, candidate gender had no impact on their evaluations..

We collect further evidence for the results of our model using administrative data from the human resources department. We begin by studying gender wage gaps in the entire set of administrative employees. In Table 9, Column 1, we show women earn about 19.8 percent less than men on average. This gap can be partially explained by job tenure and education (Column 2) and occupational sorting (Column 3), but the gap remains large and statistically significant at about 8.6 percent even after inclusion of these controls.

However, Table 10 shows that when we separate the sample by educational attainment, there is no gender wage gap among those with a BA or higher. Among those without a BA, the gender

Table 10: No gender wage gap among the highly educated

	(1) ln(Salary)	(2) ln(Salary)	(3) ln(Salary)
(β_1) Female	-0.143***	-0.188***	-0.134***
	(0.0206)	(0.0174)	(0.0232)
(β_2) BA or higher	0.584***	0.278***	0.272***
	(0.0308)	(0.0328)	(0.0314)
(β_3) Female × BA or higher	0.123***	0.196***	0.127***
	(0.0436)	(0.0382)	(0.0397)
Other controls	No	Yes	Yes
Work Unit FE	No	No	Yes
Observations	1685	1665	1665
$\beta_1 + \beta_3$	-0.02	0.008	-.007
P-val.	0.613	0.819	0.830

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

wage gap ranges from 13.4 to 18.8 percent, depending on controls (see β_1). But in each case, β_3 is positive and significant, and the sum $\beta_1 + \beta_3$ is small and statistically indistinguishable from zero, indicating no gender wage gap among the highly educated.

5 Conclusion

This paper uses a novel experimental design to study how leader gender influences the way individuals respond to leadership. We find a surprising pattern of results: while there is evidence for discrimination against female leaders when subjects have no other information about the leader, the gender gap reverses when the leader is presented as highly trained and competent. Conditional on signaling high ability, female leaders are *more* likely to be followed. Further, despite Ethiopia's poor performance on gender equity, and lower of levels of female educational attainment in general, we document a lack of discrimination in a resume evaluation experiment and no gender wage gap among the highly educated. This apparent contradiction—low levels of gender parity and education quality coupled with a lack of a gender gap among the elite—can be reconciled with a dynamic model of discrimination, in which the barriers of entry are higher for females downstream, causing for discrimination to disappear (or even reverse) at higher levels of educational attainment.

Our results in the experimental game, coupled with the results of our resume experiment and observational data, suggest that at higher levels of education and training, we may not find as much evidence of discrimination in outcomes. Importantly, however, this is not necessarily evidence of gender equality or lack of taste-based discrimination. Instead, selection into higher levels of education and training is different for women and men. If obtaining an advanced degree is harder for females, then conditional upon having an advanced degree, we may expect females to have greater ability than males. This model further suggests that as developing countries achieve gender parity in educational attainment, discrimination may begin to emerge at higher levels.

The discrimination we observe against female leadership in the absence of information is a potential explanation for why female representation in top management remains low globally despite large country-to-country variation in gender disparities in education and labor force participation. It also suggests that achieving gender parity in human capital accumulation is unlikely to solve underrepresentation of women in management.

References

African Development Bank. 2015. "Where are the women: inclusive boardrooms in Africa's top listed companies?" 1–119.

Ahmed, Shukri, and Craig Mcintosh. 2017. "the Impact of Commercial Rainfall Index Insurance: Experimental Evidence From Ethiopia."

Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, 30(2): 175.

Becker, Gary. 1957. "The Economics of Discrimination. Chicago: Univ."

Bertrand, Marianne, and Esther Duflo. 2016. "Field Experiments on Discrimination." *NBER Working Paper Series*, 22014: 110.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55(3): 789–865.

Bohren, J. Aislinn, Alex Imas, and Michael Rosenberg. 2017. "The Dynamics of Discrimination : Theory and Evidence."

Cooper, David J., and John H. Kagel. 2005. "Are two heads better than one? Team versus individual play in signaling games." *American Economic Review*, 95(3): 477–509.

Eagly, Alice H. 2013. "Women as Leaders: Leadership Style Versus Leaders' Values and Attitudes."

Guryan, Jonathan, and Kerwin Kofi Charles. 2013. "Taste-based or statistical discrimination: The economics of discrimination returns to its roots." *Economic Journal*, 123(572): 417–432.

ILO. 2016. *Women at Work: Trends 2016 - Executive Summary.* Vol. 42.

Sandberg, S, and N Scovell. 2013. *Lean in: Women, Work, and the Will to Lead. A Borzoi book*, Alfred A. Knopf.

World Bank. 2017. "World Development Indicators."

A Subject Compensation Schedule

Enumerator ID__________ Subject Number__________

Payout Schedules Provided to Subject:

Payout Schedule for Game 1: (*Show each of these as different tables at the relevant time*.)

Number of Moves – Number of Guessed Moves			Number of Moves to Solve	
0	$1.7		15	$2.00
1	$1.65		16	$1.94
2	$1.6		17	$1.88
3	$1.55		18	$1.82
4	$1.5		19	$1.76
5	$1.45		20	$1.70
6	$1.4		21	$1.64
7	$1.35		22	$1.58
8	$1.3		23	$1.52
9	$1.25		24	$1.46
10	$1.2		25	$1.40
11	$1.15		26	$1.34
12	$1.1		27	$1.28
13	$1.05		28	$1.22
14 or more, or failed to solve the puzzle.	$1		29 or more, or failed to solve the puzzle.	$1.16

Payout Schedule for Game 2:

Type A				Type B		
A's choice	Computer: In	Computer: Out		B's choice	Computer: In	Computer: Out
1	168	444		1	276	568
2	150	426		2	330	606
3	132	408		3	352	628
4	56	182		4	334	610
5	-188	-38		5	316	592

Conversion rate: 100 Points = 1 USD (e.g., 568 = 5.68)

The computer makes its decisions to try to get the maximum points possible. The computer receives points in the following way:

Computer Decides:	Type A	Type B
In	500	200
Out	250	250

Figure 5: Subject Compensation Schedule