I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10497

On the Effectiveness of Elected Male and Female Leaders and Team Coordination

Ernesto Reuben
Krisztina Timko

JANUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10497

On the Effectiveness of Elected Male and Female Leaders and Team Coordination

Ernesto Reuben
New York University Abu Dhabi and IZA

Krisztina Timko
University of Helsinki

JANUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

On the Effectiveness of Elected Male and Female Leaders and Team Coordination

We study the effect on coordination in a minimum-effort game of a leader's gender depending on whether the leader is democratically elected or is randomly-selected. Leaders use non-binding messages to try to convince followers to coordinate on the Pareto-efficient equilibrium. We find that teams with elected leaders coordinate on higher effort levels. Initially, the benefits of being elected are enjoyed solely by male leaders. However, this gender difference disappears with repeated interaction as unsuccessful male leaders are reelected more often than unsuccessful female leaders.

JEL Classification: M14, M54, J16, C92

Keywords: gender differences, leadership, democracy effect, leader effectiveness, coordination

Corresponding author:
Ernesto Reuben
New York University Abu Dhabi
PO Box 129188
Abu Dhabi
United Arab Emirates
E-mail: ereuben@nyu.edu

1. Introduction

One of the defining characteristics of effective leaders is their ability to coordinate the actions of team members through the successful communication of common goals (Bolton, Brunnermeier, and Veldkamp 2010). Historically, most leaders have been men, and to the present date, women are rare in top decision-making positions. Does this gender disparity persist because men are more effective leaders than women? In this paper, we run a laboratory experiment to investigate whether male and female leaders are equally effective in a setting where the only factor determining team success is coordination among team members. Moreover, we distinguish between teams where leaders are selected by team members and teams where they are appointed exogenously.

As many, we study team production with strong complementarities by using the minimum-effort game (Van Huyck, Battalio, and Beil 1990). In the game, team members simultaneously choose among different effort levels but team output is determined by the lowest effort. Hence, team members face a tradeoff between exerting more effort but risking that their effort is wasted because other team members do not exert effort and exerting less effort and acquiescing to a low team output. In the laboratory, play in minimum-effort games often converges to the low-effort equilibrium (for an overview see Devetag and Ortmann 2007).

A leader's role in the minimum-effort game is clear: convince other team members, the followers, that everyone will exert high effort. If followers believe others will be convinced by the leader, then it is in their best interest to exert high effort and the team will coordinate on the Pareto-dominant equilibrium. Previous work has demonstrated that giving leaders the option to send messages can substantially increase team coordination on high effort, even when a team has failed to coordinate in the past (e.g., Brandts and Cooper 2007; Sahin, Eckel, and Komai 2015; Weber et al. 2001). We use the degree to which leaders can "turnaround" a team by using non-binding pre-play messages as our measure of leader effectiveness.

We distinguish between two broad reasons why men might be more effective leaders than women in our setting. A first reason is that male leaders might behave differently than female leaders. In particular, for a given likelihood that followers will comply with a leader's message,

the fact that men are generally less risk averse (see Croson and Gneezy 2009)[1] and more overconfident (Reuben et al. 2012) suggests that male leaders are more willing to request high effort from their followers and choose high effort themselves. A second reason is that followers might hold a general belief that others are more likely to follow a message if a man sends it. If this is the case, then male leaders would be more effective even if they send the same effort request as female leaders. Given that men have historically held most of the leadership roles in society, it is reasonable to expect such beliefs by participants in the experiment.[2]

Besides gender, a leader's credibility is affected by other important factors, such as the procedure used to select the leader. In recent work, Brandts, Cooper, and Weber (2015) show that democratically-elected leaders are more effective than randomly-selected leaders, presumably because elected leaders are more legitimate.[3] We extend this line of work by studying whether male and female leaders benefit equally from being elected.[4] Since most elected leaders, such as members of parliament or heads of state are male, it is reasonable to hypothesize that elected female leaders do not benefit from the same increase in legitimacy as male leaders do.

To our knowledge, Grossman et al. (2016) is the only other study of gender differences in leadership in a coordination game.[5] Using a design similar to ours, they find that randomly-selected male leaders have a stronger impact on the behavior of followers than female leaders. Moreover, for a given impact, male leaders receive better subjective evaluations and higher monetary bonuses from their followers.

---

[1] Ertac and Gurdal (2012) demonstrate that female leaders make less risky decisions when making risky decisions for their team.

[2] For evidence that widely-held gender stereotypes affect behavior in experiments, even when they are not true in the laboratory see Reuben, Sapienza, and Zingales (2014) and Bohnet, van Geen, and Bazerman (2016).

[3] In the context of cooperation, Levy et al. (2011) show that messages from democratically-elected leaders have a bigger impact on contributions to a public good game than those of randomly-selected leaders. More generally, there is a growing body of work showing that institutions are more effective if they are democratically-chosen (see Dal Bó 2014).

[4] Even though followers do not formally vote to elect their leaders in many organizations, they are often consulted by those that do.

[5] Dufwenberg and Gneezy (2005) study differences in coordination (without leaders) depending on the fraction of men and women in a team. Other papers study gender differences in leading by example in public good games. For example, Arbak and Villeval (2013) find that men and women are equally likely to lead as long as the leaders' gender is visible.

Our study extends the work of Grossman et al. (2016) in two important ways.[6] First, we consider gender differences between elected as well as randomly-selected leaders. Second, while in Grossman et al. (2016) a team has the same leader throughout the experiment, in this paper, a team selects a leader every three periods. We think that this is an important extension because, when combined with elections, it gives teams the opportunity to reelect leaders who perform well and dispense with those who do not. Moreover, studying gender differences in reelection is particularly interesting precisely because of Grossman et al. (2016)'s finding that male leaders are evaluated more leniently than female leaders. On one hand, this bias could cause lower reelection of successful female leaders, which would exacerbate initial gender differences. On the other hand, it could also result in higher reelection of unsuccessful male leaders, which would diminish initial gender differences.

Our main findings are as follows. We find that democratically-elected leaders are more effective than randomly-selected leaders. Initially, the benefit of being elected accrues only to male leaders. However, over time, this gender difference disappears because unsuccessful female leaders are reelected at considerable lower rates than unsuccessful male leaders.

2. The experiment

Each experimental session consists of 26 periods. At the beginning of a session, participants are randomly matched into teams of five and are informed that their team's composition will not change throughout the session. In each period, every participant $i \in \{1,2,3,4,5\}$ in a team simultaneously chooses an effort level $e_i \in \{0,10,20,30,40\}$. Participant i's earnings are equal to $\pi_i = 200 - 5e_i + \min(e_1, e_2, e_3, e_4, e_5)$, where $\min(e_1, e_2, e_3, e_4, e_5)$ is the smallest effort chosen in the team. At the end of each period, participants are informed of their earnings and the team's minimum effort.

Each session is divided into two parts. Part 1 consists of periods 1 to 8 and Part 2 of periods 9 to 26. Participants know the session has two parts but are not given the specific instructions of Part 2 until they reach that part.

In Part 1, participants play the minimum effort game without a leader. Given previous evidence with these parameters (Brandts, Cooper, and Weber 2015), we expect teams will end up

[6] Another difference between our studies is that the leader makes an effort decision in our experiment. Hence, our setting is more applicable to teams where leaders perform other tasks that directly impact the team's output.

coordinating on the lowest effort level by the end of Part 1. By having teams fail to coordinate on high effort levels, we make the introduction of a leader more meaningful.

In Part 2, we introduce leaders. Specifically, every three periods, one participant in each team is selected to be the team's leader, which leaves the other team members as followers. The leader has the option to send a written message visible to all followers. The message is sent before the first period of the three-period leadership term. Leaders can write anything they wish, including nothing, except for content that could be used to identify them. Messages are non-binding in that not following a message has no direct effect on earnings. Leaders make effort decisions and face the same incentives as followers.

2.1. Leader selection

Leaders are selected with a two-step process. In the first step, each participant decides whether he or she would like to be a candidate for the leader position or not. In the second step, a leader is selected among the available candidates according to one the selection procedures described below. Candidates who are selected to be the leader receive 50 additional points. Participants who decide not to be a candidate play a lottery that pays 50 additional points with a 0.5 probability.[7]

We randomly assigned teams to one of two leader-selection procedures. In the *Random* treatment, one of the candidates is randomly assigned to be the leader, each with equal probability. In the *Election* treatment, team members elect the leader by ranking each candidate. Specifically, if there are $C > 1$ candidates, each team member i assigns a unique rank $r_i^c \in \{1, \dots, C\}$ to each candidate c, where the 1^{st} rank indicates the most preferred candidate and the C^{th} rank the least preferred. The candidate with the lowest average rank wins the election.[8] In case of a tie, the winner is chosen randomly among the tied candidates. In both *Random* and *Election*, if there are no candidates then the team has no leader for the next three periods. Moreover, if there is only one candidate then that participant automatically becomes the leader.

---

[7] Unlike Grossman et al. (2016), participants in our experiment who do not want to be a leader can opt out from becoming one. This has the disadvantage that we can only compare men and women who self-selected into the leader position. However, it has the advantage that followers know that nobody is forced to be a leader. This can matter if a leader's legitimacy varies with the leader's willingness to lead and men and women are not (believed to be) equally willing.

[8] Brandts, Cooper, and Weber (2015) use majority voting. We decided to use the Borda count because it selects leaders that are accepted by all team members rather than just a simple majority.

2.2. Gender information

To convey information about gender, we had participants chose a profile picture they identified with. This occurred after they consented to take part in the study but before they read the instructions to avoid strategic selection of profile pictures. We created 12 generic profile pictures for each gender using the profile creator website pickaface.net (see the Appendix.). All pictures had the same clothing, facial expression, face form, and eye color. We varied hair length, hair color, skin color, and did small modifications to the lips, nose, eyes, and hairstyle to match generic racial features. We use profile pictures to preserve anonymity whilst revealing gender. We opted for pictures that also contain other cues such as race and hairstyles to distract participants from discerning the purpose of the study (Zizzo 2010), which can potentially lead to intentional changes in behavior (Camerer 2015).

We displayed the profile picture of all team members during the leader-selection process and of the leader in all subsequent screens.[9] Importantly, we provide this information in both *Random* and *Election*.

2.3. Procedures

The experiment was conducted at the Vernon Smith Center of Experimental Economics (VSCEE) at Francisco Marroquín University. Participants were recruited through ORSEE (Greiner 2015) and the experiment was programed with z-Tree (Fischbacher 2007). A session lasted around one hour. We used standard experimental procedures, including random assignment of subjects to treatments, anonymity, detailed instructions with control questions, dividers between the subjects' cubicles, and monetary incentives. Earnings were expressed in points and were converted to Guatemalan quetzals at a rate of 10 quetzals per 500 points. Average earnings equaled GTQ 100.23 ($14.53). Detailed experimental procedures are available in the Appendix.

[9] Brandts, Cooper, and Weber (2015) do not provide information about gender during the leader-selectin process, but they do provide information about the candidates' performance in a trivia quiz and previous average effort level. We decided not to display this information because Brandts, Cooper, and Weber (2015) find that these variables do not influence subsequent choices.

3. Results

In total, 75 participants (15 teams) took part in *Random*, of which 36 were male and 39 were female. In *Election*, we had 70 participants (14 teams), of which 34 were male and 36 were female.

As expected, without a leader, none of the teams managed to coordinate at high effort levels. By the last period of Part 1, the minimum effort in all 29 teams was zero. Moreover, if we compare either the teams' average effort or their average earnings in Part 1 depending on the treatment and gender of the leader at the beginning of Part 2, we do not find any statistically significant differences (two-tailed Wilcoxon-Mann-Whitney U tests, $p > 0.364$ for effort and $p > 0.244$ for earnings). In other words, since all teams are in the same situation at the beginning of Part 2, any subsequent differences in behavior can be attributed to differences that occurred after leaders are introduced.

We first analyze the initial the effects of leadership, namely periods 9 to 11, in subsection 3.1. Subsequently, in subsection 3.2, we analyze the same effects in periods 12 to 26. That is, after there have been opportunities for leaders to change. Given that we have clear directional hypotheses concerning the effect of elections and the leader's gender, from here on, we report p values of one-tailed tests.

3.1. Initial effects of leadership (periods 9 to 11)

We start our analysis by looking at the decision to become a candidate. Is there a gender difference in the fraction of men and women who nominate themselves? In *Random*, 78% of men become candidates compared to 67% of women, while in *Election*, it is 79% of men and 75% of women. These gender differences are not statistically significant ($p = 0.127$ in *Random* and $p = 0.366$ in *Election*),[10] suggesting that men and women are similarly willing to become leaders. In *Election*, we can also ask the question, are male candidates more likely to be elected as leaders? We find that the answer is no, 26% of both male and female candidates became leaders.

We now turn to the effect of leaders on the teams' ability to coordinate. On the left, Figure 1 depicts the fraction of times teams coordinate on the highest effort in periods 9 to 11. On the right, it displays mean earnings during the same periods. We concentrate on these two variables as they

[10] Probit regressions with the participants' decision to become a candidate as the dependent variable and their gender as the independent variable. Standard errors clustered on teams.

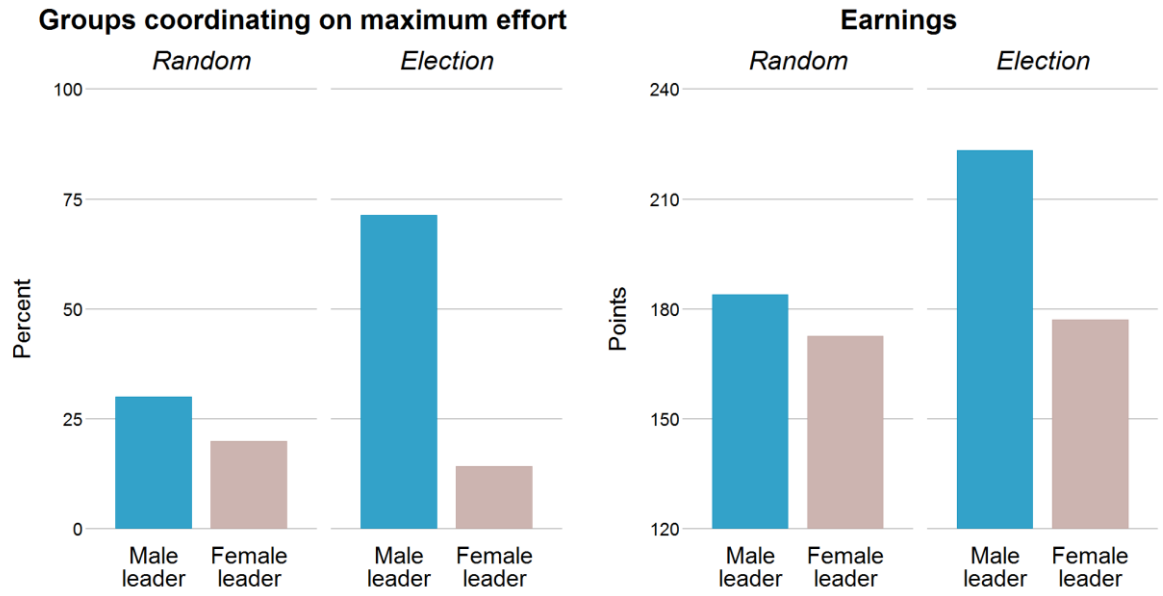


Figure 1. Fraction of teams that coordinate on the highest effort (left) and mean earnings per period (right) in periods 9 to 11 depending on the treatment and the gender of the leader.

represent the best test for a leader (coordination on the highest effort) and the teams' welfare (earnings). Additional summary statistics are available in the Appendix.

Even though leaders of both genders in both treatments manage some coordination on the highest effort, male leaders in *Election* do considerably better. The same can be said about average earnings, which are around 45 points higher with male leadership in *Election* compared to the other cases. If we test whether these differences are statistically significant, we confirm that teams in *Election* do better with male leaders than with female leaders ($p = 0.005$ for coordination, $p = 0.016$ for earnings) while teams in *Random* do not differ depending on the leader's gender ($p = 0.335$ for coordination, $p = 0.329$ for earnings). Moreover, while male leaders do better when they are elected ($p = 0.035$ for coordination, $p = 0.017$ for earnings), female leaders do not differ depending on the leader-selection procedure ($p = 0.400$ for coordination, $p = 0.436$ for earnings).[11] In other words, we replicate previous findings that leadership immediately improves coordination on high effort and that this improvement is larger with elected leaders (Brandts, Cooper, and Weber 2015). However, we find that the benefit of being elected is enjoyed exclusively by male leaders.

[11] Probit regression for coordination on the highest effort and OLS regression for earnings. Both regressions use treatment×leader's gender dummy variables and cluster standard errors on teams. All the regressions in the paper are available in the Appendix.

Table 1. Fraction of leaders who ask for the highest effort and subsequent effort choices in periods 9 to 11 depending on the treatment and the gender of the leader.

	Random		*Election*	
	Male leader	Female leader	Male leader	Female leader
% of leaders asking for the highest effort	70	60	71	43
% of teams coordinating on the highest effort				
Leader does not ask for the highest effort	0	0	0	0
Leader asks for the highest effort	43	33	100	33
% of followers choosing the highest effort				
Leader does not ask for the highest effort	0	0	4	4
Leader asks for the highest effort	63	50	100	69

Next, we take a closer look at the leaders' messages and followers' behavior to understand why teams with elected male leaders do better. We concentrate on one specific type of message: explicitly asking followers to choose the highest effort. The benefit of looking at this type of message is that it is easily coded and it has been shown by Brandts, Cooper, and Weber (2015) to be crucial for team coordination.

The first row of Table 1 shows that fraction of leaders who ask for the highest effort ranges from 71% of male leaders in *Election* to 43% of female leaders in *Election*. Even though this is a noticeable difference in the leaders' behavior, it is not statistically significant.[12] The subsequent rows of Table 1 show the fraction of teams that coordinate on the highest effort and the fraction of followers who choose the highest effort depending on the leader's message. Clearly, if a leader does not ask for the highest effort then followers do not choose high effort irrespective of the treatment or the leader's gender. By contrast, if a leader does ask for the highest effort, the followers' reaction varies significantly. Specifically, all the followers of elected male leaders follow the leader's request whereas a considerably smaller fraction does so if the leader is female or is randomly appointed.[13] We summarize these findings as our first result.

[12] Male vs. female leaders in *Election* ($p = 0.134$) and in *Random* ($p = 0.354$). *Election* vs. *Random* among male ($p = 0.475$) and female ($p = 0.721$) leaders. Probit regression with treatment×leader's gender dummy variables and standard errors clustered on teams (see the Appendix).

[13] Male vs. female leaders in *Election* ($p = 0.008$ for coordination, $p = 0.015$ for effort) and in *Random* ($p = 0.389$, for coordination, $p = 0.296$ for effort). *Election* vs. *Random* among male ($p = 0.001$, for coordination, $p = 0.002$ for effort) and female ($p = 0.500$, for coordination, $p = 0.219$ for effort) leaders. Since followers react only when the leader asks for the highest effort, we use Probit regressions with sample selection based on the type of message sent, treatment×leader's gender dummy variables, and standard errors clustered on teams (see the Appendix).

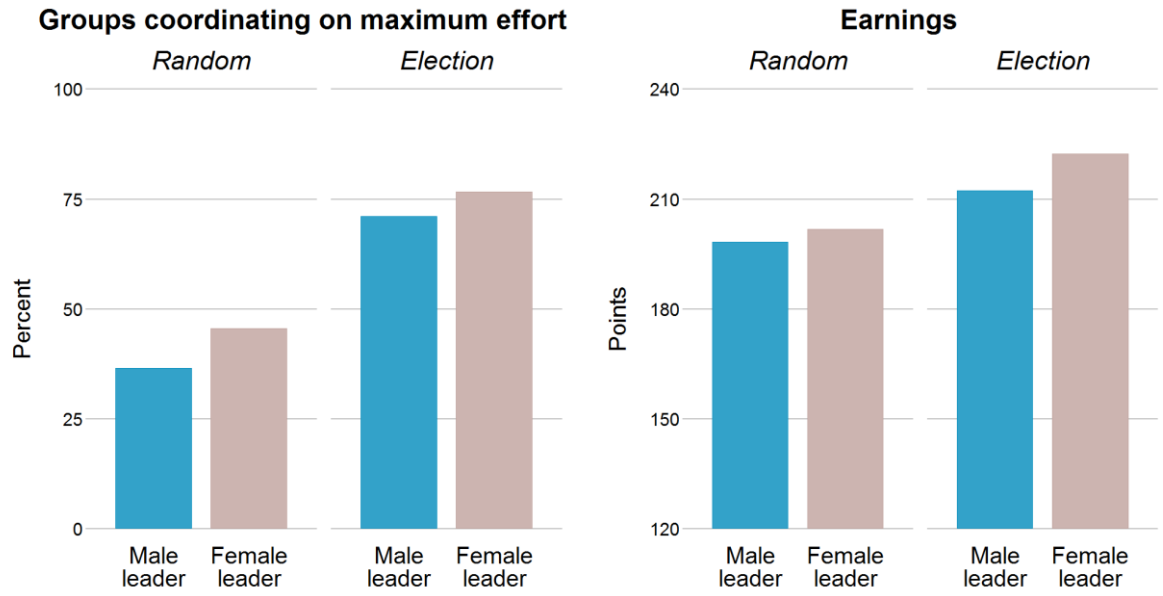


Figure 2. Fraction of teams that coordinate on the highest effort (left) and mean earnings per period (right) in periods 12 to 26 depending on the treatment and the gender of the leader.

Result 1: *Teams that elect a male leader initially do better than teams that elect a female leader and teams with randomly-chosen leaders of either gender. This is mostly due to the behavior of followers, who are more likely to follow the requests for high effort of elected male leaders.*

3.2. Effects of leadership after reselection (periods 12 to 26)

Next, we evaluate whether the advantage of elected male leaders remains after leader reselection has taken place. Figure 2 and Table 2 present the same statistics as Figure 1 and Table 1 but for periods 12 to 26.

Over time, the advantage of elected male leaders over randomly-selected leaders decreases but still exists. By contrast, the performance of elected female leaders improves to the point that it equals that of their male counterparts (see Figure 2). This results in a difference between elected and randomly-appointed leaders but no difference between male and female leaders.[14] Table 2 reveals why this is the case. Both male and female leaders almost always request the highest effort in *Election* but not in *Random*.[15] Moreover, over time, the differences in the reaction of followers to high-effort requests disappear.

[14] *Election* vs. *Random* among male (coordination: $p = 0.018$; earnings: $p = 0.141$) and female (coordination: $p = 0.055$; earnings: $p = 0.077$) leaders. Male vs. female leaders in *Election* (coordination: $p = 0.711$; earnings: $p = 0.866$) and in *Random* (coordination $p = 0.746$; earnings: $p = 0.640$ in *Random*). Same regressions as footnote 11.

[15] *Election* vs. *Random* among male ($p = 0.017$) and female ($p = 0.085$) leaders. Same regressions as footnote 12.

Table 2. Fraction of leaders who ask for the highest effort and subsequent effort choices in periods 12 to 26 depending on the treatment and the gender of the leader.

	Random		*Election*	
	Male leader	Female leader	Male leader	Female leader
% of leaders asking for the highest effort	69	70	92	90
% of teams coordinating on the highest effort				
Leader does not ask for the highest effort	0	0	0	0
Leader asks for the highest effort	53	65	77	85
% of followers choosing the highest effort				
Leader does not ask for the highest effort	0	0	0	0
Leader asks for the highest effort	62	74	84	89

To conclude our analysis, we take a closer look at behavior in elections to understand why female leaders catch up with male leaders. Specifically, we look at whether the leaders' gender impacts the fraction of followers who challenge the leader by becoming candidates and the likelihood that the leader is reelected. Table 3 presents both statistics depending on whether the team coordinated on the highest effort during the leader's three-period term. Not surprisingly, leaders of successful teams are not challenged frequently and are likely to be reelected. For these teams, we do not observe significant gender differences ($p > 0.197$). By contrast, we do find an important gender difference among unsuccessful teams. While unsuccessful male and female leaders are challenged at similar rates ($p = 0.700$), unsuccessful male leaders have a significantly higher chance of being reelected ($p = 0.041$). These findings are stated as our second result.

Result 2: *Over time, although elected leaders do better than randomly-appointed leaders, male and female leaders perform equally well. An important contributing factor is that unsuccessful female leaders are reelected less often than unsuccessful male leaders, which reduces the gender difference in performance of the remaining leaders.*

4. Conclusions

We study the gender differences in the effectiveness of non-binding messages sent by team leaders in a minimum-effort game depending on whether a leader is elected or is randomly selected.

Our findings are partly consistent with those of Grossman et al. (2016). We find a gender difference in the effectiveness of leaders but only for elected leaders and not for randomly-selected ones. Unfortunately, there are numerous design differences between the two studies to

Table 3. Likelihood of followers challenging and of the leader being reelected in periods 12 to 26 depending on the leader's gender and team coordination on the highest effort.

	Leader was	
	Male	Female
Team did not coordinate on high effort in the last three periods		
Fraction of followers who become candidates	48	50
Fraction of leaders who are reelected	67	20
Team coordinated on high effort in the last three periods		
Fraction of followers who become candidates	35	22
Fraction of leaders who are reelected	76	85

pinpoint the precise reason for this disparity. For example, in our study, team members decide whether to become a candidate, leaders make an effort decision, messages are written instead of spoken, and leaders change every three periods. However, our findings do call for caution as they suggest that gender differences in leadership effectiveness are not general and depend on the precise context in which they are studied.

Like Grossman et al. (2016), we find compelling evidence that male and female leaders are evaluated differently for a given team performance. In our study, this bias results in weaker selection pressure among male leaders. Interestingly, lenience towards unsuccessful male leaders contributes to the disappearance of the initial gender difference in effectiveness. This result is a useful reminder that observing gender parity in an organization might in fact be due to gender disparities in evaluations. It also suggests that policies such as gender quotas might not only increase female representation but also the competence of male leaders (Baltrunaite et al. 2014).

References

Arbak, Emrah, and Marie-Claire Villeval. 2013. "Voluntary Leadership: Motivation and Influence." *Social Choice and Welfare* 40(3): 635–62.

Baltrunaite, Audinga, Piera Bello, Alessandra Casarico, and Paola Profeta. 2014. "Gender Quotas and the Quality of Politicians." *Journal of Public Economics* 118: 62–74.

Bohnet, Iris, Alexandra van Geen, and Max Bazerman. 2016. "When Performance Trumps Gender Bias: Joint vs. Separate Evaluation." *Management Science* 62(5): 1225–34.

Bolton, Patrick, Markus K. Brunnermeier, and Laura Veldkamp. 2010. "Economists' Perspectives on Leadership." In *Handbook of Leadership Theory and Practice*, eds. Nitin Nohria and Rakesh Khurana. Boston, MA: Harvard Business School Press, 239–64.

Brandts, Jordi, and David J Cooper. 2006. "A Change Would Do You Good .... An Experimental Study on How to Overcome Coordination Failure in Organizations." *American Economic Review* 96(3): 669–93.

Brandts, Jordi, and David J. Cooper. 2007. "It's What You Say, Not What You Pay: An Experimental Study of Manager–Employee Relationships in Overcoming Coordination Failure." *Journal of the European Economic Association* 5(6): 1223–68.

Brandts, Jordi, David J. Cooper, and Roberto A. Weber. 2015. "Legitimacy, Communication, and Leadership in the Turnaround Game." *Management Science* 61(11): 2627–45.

Camerer, Colin F. 2015. "The Promise and Success of Lab–Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List." In *Handbook of Experimental Economic Methodology*, Oxford University Press, 249–95.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2): 448–74.

Dal Bó, Pedro. 2014. "Experimental Evidence on the Workings of Democratic Institutions." In *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*, eds. Sebastian Galiani and Itai Sened. New York, NY: Cambridge University Press, 266–88.

Devetag, Giovanna, and Andreas Ortmann. 2007. "When and Why? A Critical Survey on Coordination Failure in the Laboratory." *Experimental Economics* 10(3): 331–44.

Dufwenberg, Martin, and Uri Gneezy. 2005. "Gender & Coordination." In *Experimental Business Research*, eds. Rami Zwick and Amnon Rapoport. Boston, MA: Springer US, 253–62.

Ertac, Seda, and Mehmet Y. Gurdal. 2012. "Deciding to Decide: Gender, Leadership and Risk-Taking in Groups." *Journal of Economic Behavior & Organization* 83(1): 24–30.

Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10(2): 171–78.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1(1): 114–25.

Grossman, Philip J, Catherine C Eckel, Mana Komai, and Wei Zhan. 2016. *It Pays to Be a Man: Rewards for Leaders in a Coordination Game*.

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil. 1990. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review* 80(1): 234–48.

Levy, David M. et al. 2011. "Leadership, Cheap Talk and Really Cheap Talk." *Journal of Economic Behavior & Organization* 77(1): 40–52.

Reuben, Ernesto, Pedro Rey-Biel, Paola Sapienza, and Luigi Zingales. 2012. "The Emergence of Male Leadership in Competitive Environments." *Journal of Economic Behavior & Organization* 83(1): 111–17.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2014. "How Stereotypes Impair Women's Careers in Science." *Proceedings of the National Academy of Sciences* 111(12): 4403–8.

Sahin, Selhan Garip, Catherine Eckel, and Mana Komai. 2015. "An Experimental Study of Leadership Institutions in Collective Action Games." *Journal of the Economic Science Association* 1(1): 100–113.

Weber, Roberto, Colin Camerer, Yuval Rottenstreich, and Marc Knez. 2001. "The Illusion of Leadership: Misattribution of Cause in Coordination Games." *Organization Science* 12(5): 582–98.

Zizzo, Daniel John. 2010. "Experimenter Demand Effects in Economic Experiments." *Experimental Economics* 13(1): 75–98.

Appendix

The first section of the Appendix contains more detail about the experiment's procedures, including a detailed timeline, a sample of the instructions, and screenshots of the computer program. The second section contains descriptive statistics as well as the regressions reported in the main body of the paper.

A.1. Detailed experimental procedures

After their arrival to the laboratory, participants were assigned randomly to seats. Before period 1, and before reading the instructions, everybody answered a short general questionnaire about gender, race, age, years of study, and major field of studies. Next, participants had to choose a profile picture. Figure A1 contains the 24 profile pictures.

We had separate instructions for Part 1 and Part 2, and participants read the instructions only prior to each part. To facilitate calculations for the participants, we handed out printed versions of the instructions for Part 1, which contained a table showing how earnings were determined in each period. The same table applied in Part 2. Instructions were displayed on the computer screens and were read aloud by the experimenter. After reading the instructions for Part 1, participants completed a payoff quiz to check whether everybody understood the game's payoff structure. Participants completed a questionnaire about the volunteering and selection process after reading the instructions for Part 2. Instructions and screenshots can be found below.

The game was described using a workplace context to be in line with earlier papers, ease comprehension of the task, and enrich the wording and analysis of the free form messages (see Brandts and Cooper 2006). As in Brandts, Cooper, and Weber (2015), individual team members were referred to as "employee", and they were told that they were working for a "firm". The leader

Figure A1. Available profile pictures to female (left) and male (right) participants

was called the "manager". Following Brandts, Cooper, and Weber (2015), we did not use the term "effort" because of its strong connotation. Instead, we asked participants to think of each period as a "workweek" lasting 40 hours and choose how many hours to devote to the firm's "bonus project".

Leaders could enter their messages into a chat box, and they could either click on a button to send the message or click on a button labeled as "Send no suggestion". The message, along with the leader's profile picture and ID number, was displayed on all screens throughout the three periods of a leadership term. We used the profile pictures along with unique ID numbers to display candidates and selected leaders to all team members. Participants knew from the instructions of Part 2 that the profile pictures will be displayed.

In Part 2 we elicited some of the participants' beliefs. After they decided whether to be candidate or not, we elicited their belief concerning the number of other candidates by asking "Out of the four other participants in your firm, how many will run for the Manager position?" After participants made their effort choice, we elicited their belief concerning the number of other team members who will follow the leader's message by asking "Out of the four other participants in your firm, how many will follow the Manager's suggestion?" These belief questions were only asked in first period of each leadership term. Table A1 summarizes the sequence of events in the experiment.

At the end of each period, participants saw their effort choice, the team minimum effort, their earnings in that period, and their accumulated earnings. Participants could not observe individual effort choices. At the end of the experiment participants were shown their earnings in points and

Table A1. Timeline of the experiment

Timeline	Activity
Before period 1	▪ Indicate demographic characteristics ▪ Choose a profile picture ▪ Instructions for Part 1 ▪ Control questions for Part 1
Every period in periods 1 to 8	▪ Effort choice ▪ Feedback screen
Before period 9	▪ Instructions for Part 2 ▪ Control questions for Part 2
Every 3 periods in periods 9 to 26	▪ Candidacy choice ▪ Belief question about number of candidates ▪ Leader selection through rankings (*Election*) or a random draw (*Random*) ▪ Selection and lottery results ▪ Leader sends message to followers
Every period in periods 9 to 26	▪ Effort choice
Every 3 periods in periods 9 to 26	▪ Belief question about number of followers
Every period in periods 9 to 26	▪ Feedback screen

Guatemalan quetzals, including any leader and lottery bonuses. Participants were thanked and paid individually for their participation.

A.1.1. Sample instructions

Thank you for participating in this session. You are participating in a study on economic decision making and will be asked to make a number of decisions. Please read these instructions carefully as they describe how you can earn money.

All the interaction between you and other participants will take place through the computers. Please do not talk or communicate in any other way with other participants. If you have a question, raise your hand and one of us will help you. The study is *anonymous*: that is, your identity will not be revealed to others and the identity of others will not be revealed to you.

During the study your earnings will be expressed in points. Upon completion of the session, your accumulated earnings will be converted from points to quetzals at a rate of *10 quetzals per 500 points*. You will be paid these converted earnings in cash. The study is divided into *two parts*. Your earnings will equal the sum of earnings from each part. You will be randomly assigned to a firm of *five* participants. You will be grouped with the same five participants throughout Part 1 and Part 2. Part 1 consists of *8 periods* and Part 2 consists of *18 periods*. You will read the instructions for Part 1. You will receive the instructions for Part 2 once Part 1 has been completed.

Part 1

You are one out of five employees in a firm. Each period can be thought of as a workweek. Each employee spends 40 hours per week at their firm. In each period, every employee will be asked to choose how many hours to devote to the firm's bonus project. The available choices are 0 hours, 10 hours, 20 hours, 30 hours, and 40 hours.

The earnings for an employee are determined in each period by how many hours that employee spends on the bonus project, and the minimum number of hours employees in his or her firm spend on the bonus project. Specifically, the employee's earnings are reduced by 5 points per hour that he or she spends on the bonus project. In addition, the employee also receives a bonus equal to the *minimum* number of hours *any* employee in his or her firm spends on the bonus project multiplied by 6 points. Each employee also gets a flat payoff of 200 points in each period. In other words, your earnings are given by the formula below:

200 – 5 × your hours in bonus project + 6 × minimum hours in bonus project by any employee

To facilitate your calculations, the following Earnings Table shows how your earnings depend on your choice and the minimum choice in your firm.

EARNINGS TABLE

		Minimum number of hours chosen in the firm				
		40	**30**	**20**	**10**	**0**
Your hours	**40**	240	180	120	60	0
	30		230	170	110	50
	20			220	160	100
	10				210	150
	0					200

Your earnings in each period are found by looking across from the number of hours you chose on the left-hand side and down from the minimum number of hours chosen in the firm by any employee. For example, suppose you spend 10 hours on the bonus project. Suppose the other four employees in the firm spend 20, 30, 40 and 40 hours. The minimum hours spent on the bonus project is 10 hours. Then your payoff equals: 200 - 5×10 + 6×10 = 210 points.

At the end of each period you will receive a summary of what happened in the period including the number of hours you spent on the bonus project, the minimum number of hours chosen in the firm, your payoff for the latest period, and your accumulated payoffs for the current part. The

computer also provides a summary of this information for preceding periods. At no point in time will we identify the identity of any employees in the firm. In other words, the actions you take will remain confidential. To ensure your understanding of these instructions, click the "READY" button and answer the questions that will appear on your screen.

Part 2

Part 1 has ended. Read the instructions for Part 2 and click on READY once you are done. Part 2 is similar to Part 1. Now there will be *18 periods* in Part 2. In each period, every employee will choose how many *hours* to devote to the firm's bonus project. Available choices are 0, 10, 20, 30 and 40 hours. The number of hours you choose and the *minimum number of hours* chosen in the firm will determine your earnings in that period. The Earnings Table is the same as in Part 1. Finally, your firm's composition has not changed. In other words, in Part 2 you will interact with the same firm of *five* people as in Part 1.

The difference between Part 1 and Part 2 is that in *every 3 periods* there will be a selection and a message stage.

Selection stage

In the selection stage one person in your firm will be selected to be the *Manager*. The selection stage goes as follows:

- First, all employees decide whether they wish to run for the manager position. Employees who run are referred to as candidates.
- Thereafter, all employees (both candidates and non-candidates) vote to elect a manager. During the vote, employees can identify the candidates by a randomly assigned id (a number between 1 and 5) and their chosen profile picture.
- Employees vote by ranking the candidates from most preferred to least preferred. For example, if there are three candidates, each employee has to assign one candidate the rank of 1 (most preferred), another candidate the rank of 2 (second most preferred), and the remaining candidate to the rank of 3 (least preferred). Note that candidates must also rank themselves when they vote.
- The candidate with the best average rank wins the election and becomes the manager. In case of a tie, the winner will be chosen randomly among the tied candidates.
- If none of the employees runs, then there is no election and the firm will not have a manager.

- If only one employee runs, then that employee automatically wins the election and becomes the manager.

Earnings from the selection stage

You can earn additional money in the selection stage. Additional earnings will be added to your total earnings at the end of the experiment.

- Employees who run for the manager position will earn *50 points if they win* the election and *0 points if they lose* the election.
- Employees who do not to run for the manager position will earn *50 points* with 50% probability and *0 points* otherwise. Whether you earn 50 points or 0 points is determined randomly by the computer.

You will be informed of the selection-stage earnings immediately after the election.

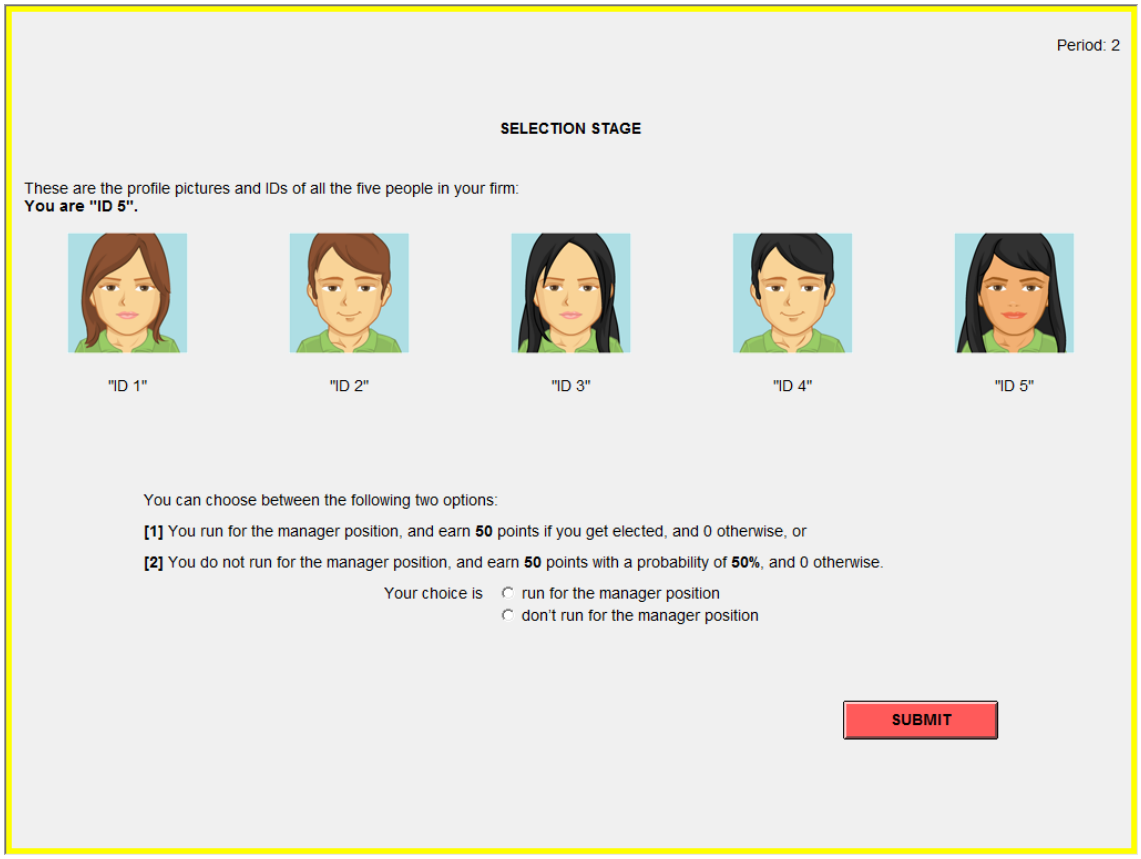
Message stage

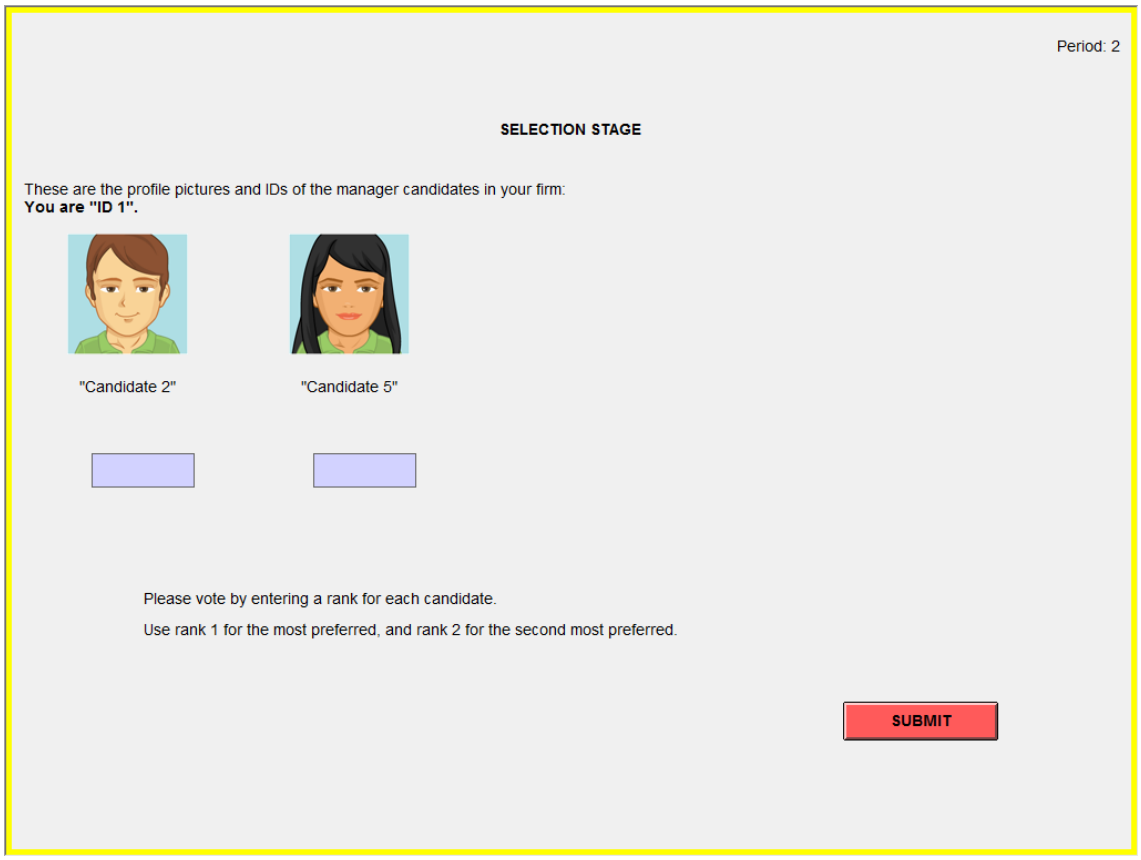
After the vote, you will be informed whether you have been assigned the role of the Manager or the role of an Employee. The Manager's profile picture will be displayed on the computer screen. In the message stage, the *Manager* will be able to send a *written suggestion* to all employees, or alternatively, he/she can decide not to send any suggestion. The suggestion cannot contain information that can be used to identify the Manager, such as a name, nickname, or any other identifying feature like clothing, or the desk number. Other than these restrictions, the Manager may write anything that he/she wishes. After the message stage, every employee will see the Manager's suggestion.

Subsequently, employees and managers play for 3 periods. In each period, each employee and the Manager enter the *number of hours* they wish to choose. Note that the suggestion does not commit you to any particular choice. That is, neither the Manager nor the other employees are required to choose the number of hours indicated in the suggestion.

After 3 periods, the manager reverts to being employee and there will be a new selection and message stage. Employees will make new decisions about candidacy and voting. Note that ids are fixed throughout Part 2.

A.1.2. Screenshots





You are the Manager Period: 2

SELECTION RESULTS

Candidate "ID 5" was elected as Manager in your firm.

You have been assigned to the role of **MANAGER**

You ran for the manager position and you **won**. You earned **50** points.

CONTINUE

You are the Manager Period: 2

MESSAGE STAGE

Write your suggestion to all employees in the box below.

Remember to hit first the [Enter] key before you click on FINISHED SENDING SUGGESTIONS!
If you wish to send "No suggestion", click on SEND NO SUGGESTION and confirm with YES.

SEND NO SUGGESTION FINISHED SENDING SUGGESTIONS

You are an Employee

Period: 2

DECISION STAGE

Manager "ID 5".

The suggestion from the Manager to all Employees is:

please choose 40

Please enter your hours

- 40
- 30
- 20
- 10
- 0

SUBMIT

You are an Employee

Period: 2

RESULTS

Manager "ID 5".

The suggestion from the Manager to all Employees was:

please choose 40

Your hours were: 40

The minimum number of hours was: 40

Your earnings (in points) in this period are: 240

So far, your earnings (in points) in Part 2 are: 240

CONTINUE

Table A2. Descriptive statistics

		First leader (periods 9-11)				Later leaders (periods 12-26)			
Treatment		*Random*		*Election*		*Random*		*Election*	
Leader's gender		Male	Female	Male	Female	Male	Female	Male	Female
Team's minimum effort	mean	15	8	29	15	16	19	28	31
	s.d.	(18)	(17)	(19)	(16)	(19)	(20)	(18)	(17)
% coordinating on the highest effort	mean	30	20	71	14	37	46	71	77
	s.d.	(47)	(41)	(46)	(36)	(48)	(50)	(46)	(43)
Individual effort	mean	21	15	30	22	19	23	32	32
	s.d.	(19)	(19)	(17)	(17)	(19)	(19)	(16)	(16)
Earnings	mean	184	173	223	177	198	202	212	222
	s.d.	(72)	(77)	(38)	(76)	(57)	(59)	(63)	(48)

A.2. Additional statistical analysis

Table A2 contains descriptive statistics of the teams' minimum effort, the percentage of teams coordinating on the highest effort, the effort of individual participants, and individual earnings depending on the treatment, the gender of the leader, and whether the current leadership term is the first term with a leader or not.

Table A3 presents estimates from regressions testing the effect of a leader's gender and the leader's selection procedure on team coordination, earnings, and the likelihood that a leader requests the highest effort. In column I, the dependent variable, in each period, equals one if a team coordinates on the highest effort and zero otherwise. It presents marginal effects (in percentages) from a probit regression. In column II, the dependent variable equals the participants' earnings in each period. It presents coefficients from an OLS regression. In column III, the dependent variable, in each period, equals one if a leader asks the followers for the highest effort and zero otherwise. It presents marginal effects (in percentages) from a probit regression. In all regressions, we use as independent variables the interaction of treatment (*Election* or *Random*), the gender of the leader in a period (male or female), and whether it is the team's first leadership term or not (periods 9 to 11 or periods 12 to 26). In addition, we also add dummy variables to identify periods in which teams had no leader because nobody chose to become a candidate, which occurred in 3 percent of all periods in *Random* and 2 percent in *Election*. In all regressions, we cluster standard errors on teams. The regressions in columns I and II include data from periods 9 to 26, while the regression in column II includes data from every third period in

Table A3. Team coordination, earnings, and leaders' message choice depending on the leader-selection procedure and the leader's gender

Note: (I) Marginal effects in percentages from a probit regression with team coordination on the highest effort as the dependent variable. (II) Coefficients from an OLS regression with individual earnings as the dependent variable. (III) Marginal effects in percentages from a probit regression with a request for the highest effort as the dependent variable. Standard errors corrected for clustering on teams in parenthesis. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

	I	II	III
Female leader	−57***	−46**	−29
	(22)	(21)	(26)
Random	−41*	−39**	−1
	(23)	(18)	(23)
Random × Female leader	−51**	−51**	−11
	(25)	(23)	(28)
Periods 12 to 26	0	−11	21
	(13)	(9)	(17)
Periods 12 to 26 × Female leader	5	−1	19
	(16)	(8)	(19)
Periods 12 to 26 × *Random*	−35	−25*	−2
	(22)	(13)	(20)
Periods 12 to 26 × *Random* × Female leader	−26	−22	−1
	(23)	(15)	(21)
Constant	71***	223***	71***
	(17)	(10)	(17)
Obs.	522	2610	174
Clusters	29	29	29
R^2 / Pseudo R^2	12	5	9

periods 9 to 26 (when messages could be sent). Note that the omitted category is teams with a male leader in periods 9 to 12 in *Election*.

Table A4 presents marginal effects (in percentages) from probit models with sample selection testing the effect of a leader's gender and the leader's selection procedure on team coordination and the followers' individual effort. The regressions correct for the observation that followers react positively only when leaders request the highest effort, and leader requests can be affected by their gender and the selection procedure. In column I, the dependent variable, in each period, equals one if a team coordinates on the highest effort and zero otherwise. In column II, the dependent variable, in each period, equals one if a follower chooses the highest effort and zero otherwise. In both regressions, the first stage consists of a probit regression with a dependent variable that, in each period, equals one if a leader asks for the highest effort. In both regressions, we use as independent variables the interaction of treatment (*Election* or *Random*) and the

Table A4. Team coordination and followers' effort choice given that the leader requests the highest effort depending on the leader-selection procedure and the leader's gender

Note: Conditional marginal effects in percentages from probit regressions with sample selection with (I) team coordination on the highest effort as the dependent variable and (II) the followers' choice of the highest effort as the dependent variable. Standard errors corrected for clustering on teams in parenthesis. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

	I	**II**
Female leader	−67**	−30**
	(28)	(14)
Random	−57***	−37***
	(19)	(13)
Random × Female leader	−67**	−50**
	(28)	(21)
Constant	100***	100***
	(0)	(0)
Obs.	87	381
Uncensored obs.	54	216
Clusters	29	29
Wald χ^2	230	351

gender of the leader (male or female). In both regressions, we cluster standard errors on teams and include data from periods 9 to 12. Note that the omitted category is teams with a male leader in *Election*.

Table A5 presents marginal effects (in percentages) from probit models testing the effect of a leader's gender on the likelihood that followers challenge the leader in the next election and the likelihood that leaders are reelected. In column I, the dependent variable equals one if a follower decides to become a candidate in the current election and zero otherwise. In column II, the dependent variable equals one if a candidate wins the current election and zero otherwise. In both regressions, we use, as independent variables, dummy variables to identify the gender of the leader in the previous three periods and whether the team was successful or unsuccessful (i.e., whether the team coordinated on the highest effort or not). Hence, we include a dummy variable that equals one if the previous leader was female and the team was unsuccessful, another dummy variable that equals one if the previous leader was male and the team was successful, and a last dummy variable that equals one if the previous leader was female and the team was successful. Note that this means that the omitted category is a previous male leader and an unsuccessful team. In both regressions, we cluster standard errors on teams and include data from every election in periods 12 to 26 (the election in period 9 is not used since there is no previous leader).

Table A5. Likelihood of followers challenging and of the leader being reelected depending on the leader's gender and team coordination on the highest effort

Note: Marginal effects in percentages from probit regressions with (I) the followers' decision to become a candidate as the dependent variable and (II) leaders' reelection conditional on being a candidate as the dependent variable. Standard errors corrected for clustering on teams in parenthesis. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

	I	**II**
Unsuccessful female leader	2	–47**
	(6)	(23)
Successful male leader	–13*	10
	(8)	(23)
Successful female leader	–25**	18**
	(11)	(25)
Constant	48***	67***
	(6)	(22)
Obs.	282	46
Clusters	14	13
Wald χ^2	6	17

The regression in column I includes decisions only from followers while the regression in column II includes election outcomes only for leaders (i.e., to determine reelection).