Online Appendices for the paper:
On the Effectiveness of Elected Male and Female Leaders and Team Coordination

Ernesto Reuben

New York University Abu Dhabi, e-mail: ereuben@nyu.edu

Krisztina Timko

University of Helsinki, e-mail: krisztina.timko@gmail.com

ABSTRACT

This document contains the supplementary materials for Reuben and Timko (2018). The first section 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. The third section includes additional findings not reported in the paper due to space constraints.

A. Detailed experimental procedures

The experiment was conducted at the Vernon Smith Center of Experimental Economics (VSCEE) at Francisco Marroquín University. Seven sessions where conducted in May 2015 (four for *Random* and three for *Election*) and two more sessions where conducted in May 2017 (both for *Election*) at the request of the papers' reviewers. After their arrival to the laboratory, participants were randomly assigned to seats. 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. The profile pictures were chosen to represent the racial/ethnic composition of the subject pool.

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

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

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 of the free-form messages (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 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 questions were asked in the first period of each leadership term. Table A1 summarizes the experiment's sequence of events.

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

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

A.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 your earnings in cash. The study has *two parts*. Your earnings today 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 below. 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.

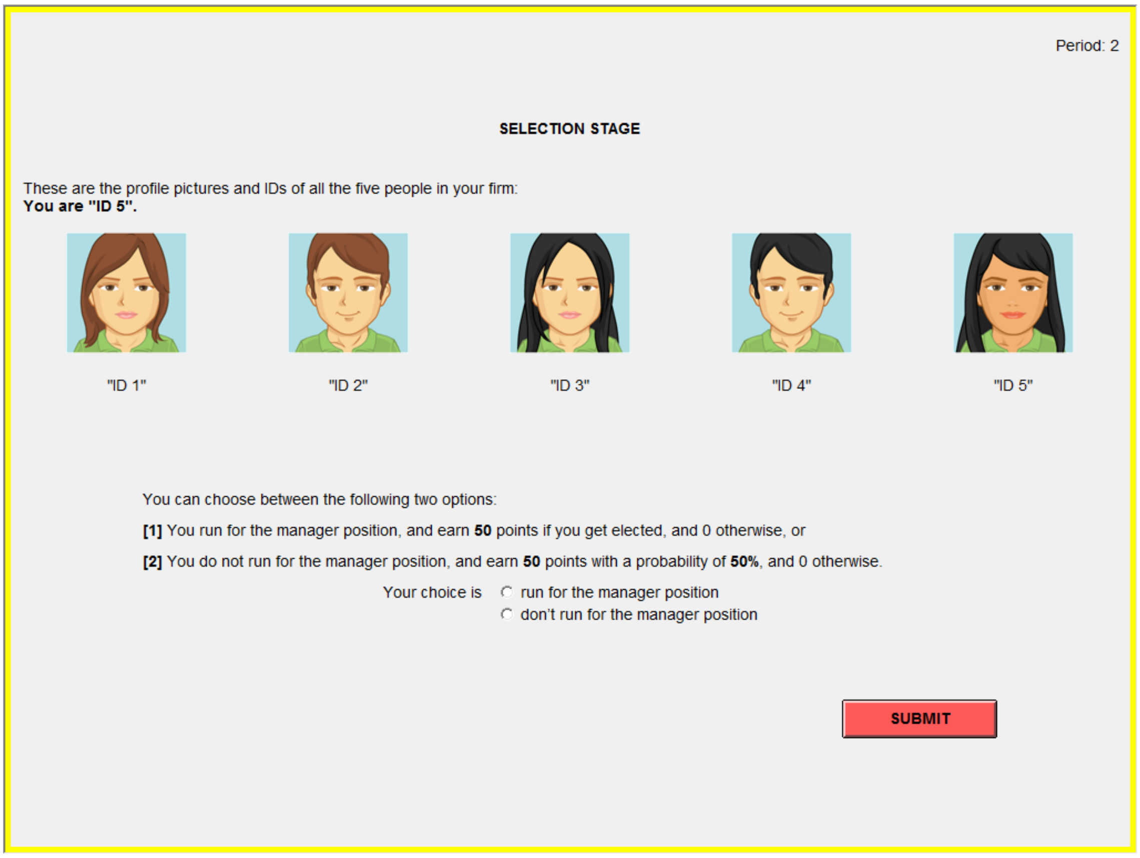
Message stage

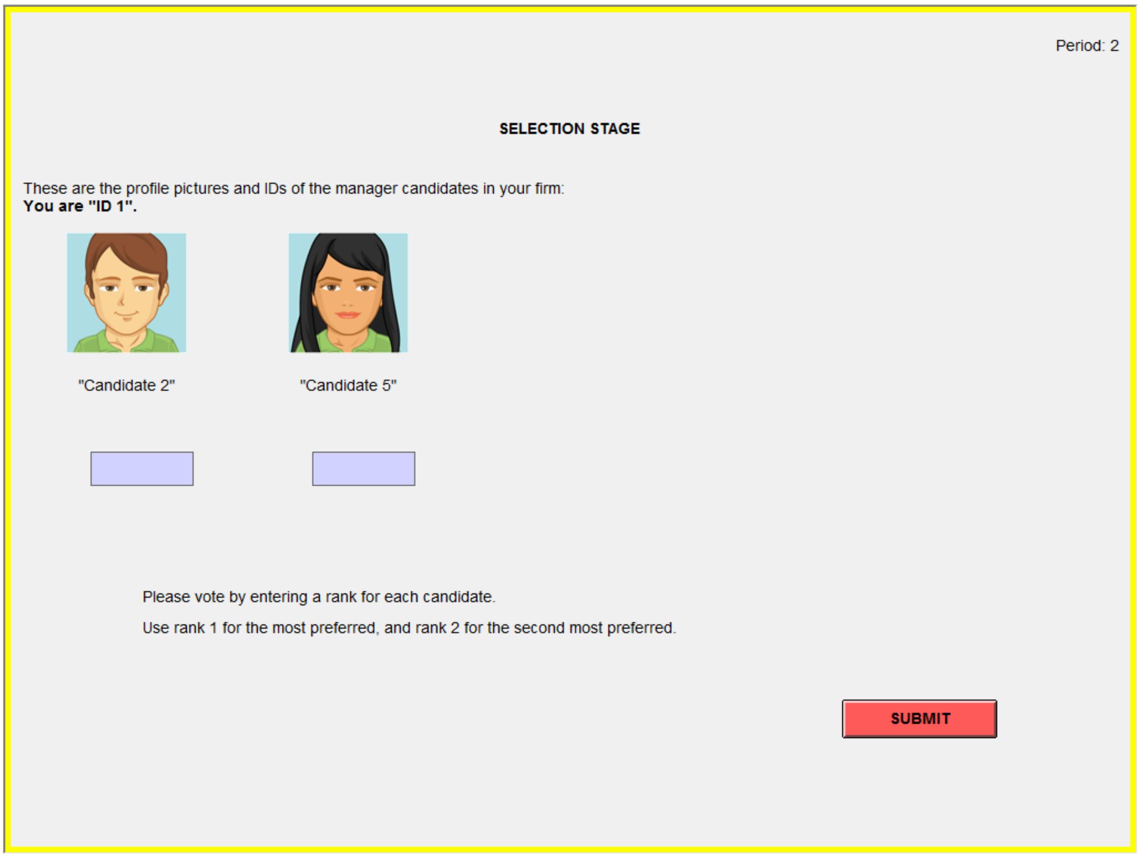
After the vote, you will be informed whether you have been assigned the role of the Manager or the role of an Employee. The Manager's profile picture will be displayed on the computer screen. In the message stage, the *Manager* will be able to send a *written suggestion* to all employees, or alternatively, 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.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

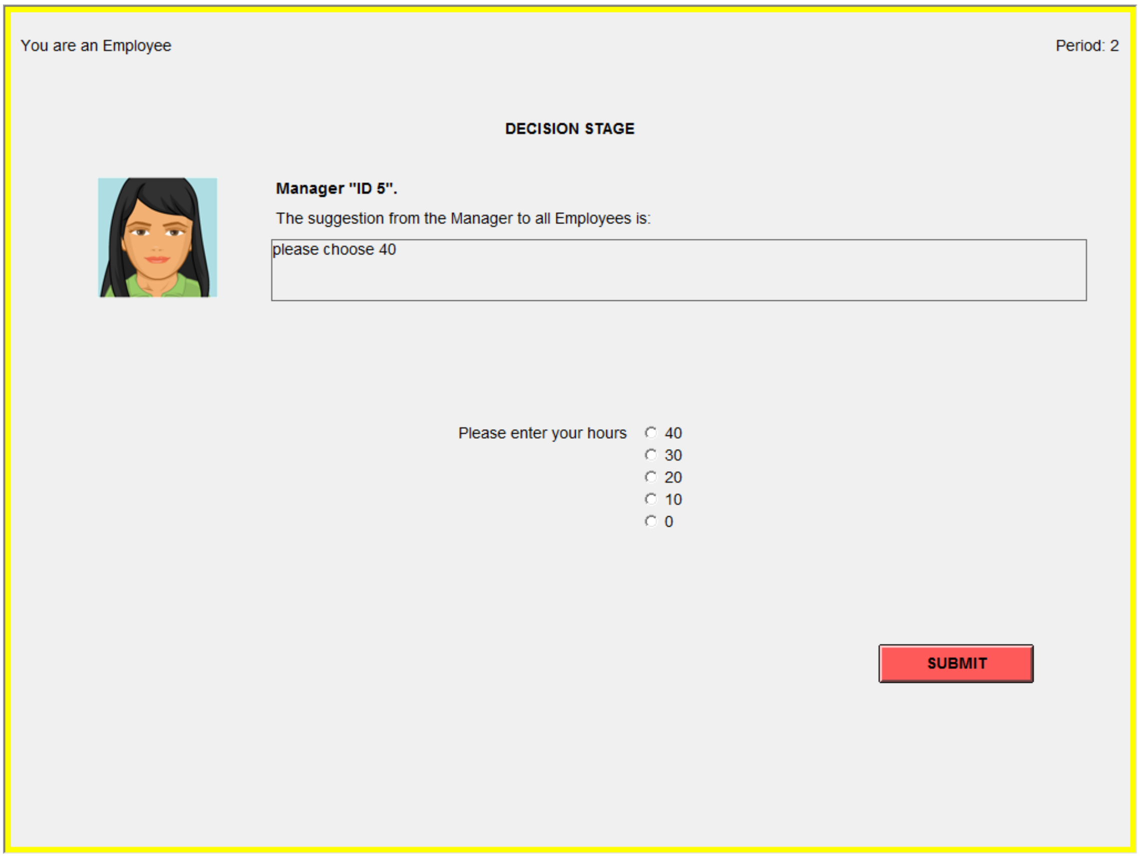
You are the Manager Period: 2

MESSAGE STAGE

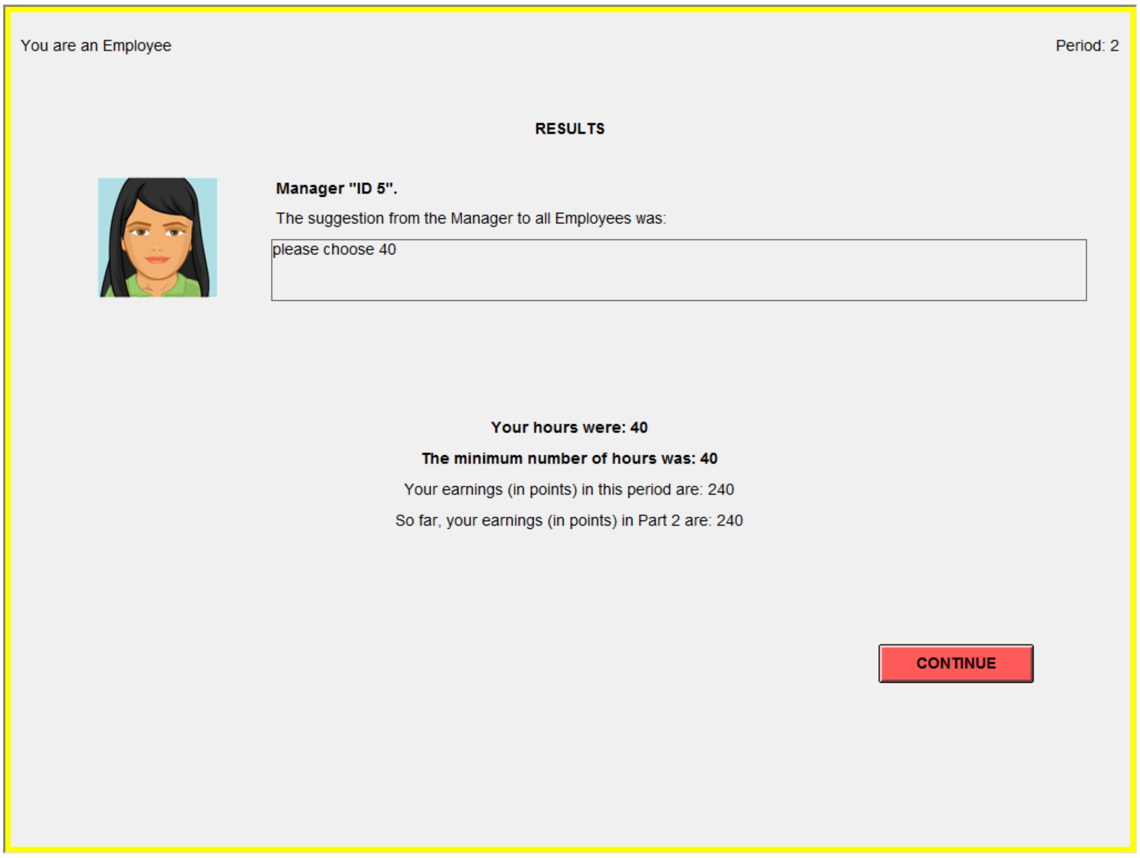
Write your suggestion to all employees in the box below.

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. Effort and earnings in Part 1 depending on the leader-selection procedure and the leader's gender in period 9

Note: (I) Marginal effects in percentages of the probability of exerting an effort of zero based on an ordered probit regression with individual effort as the dependent variable. (II) Coefficients from an OLS regression with individual earnings 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
Election	1	1
	(7)	(6)
Random × Female leader	−7	−7**
	(8)	(7)
Election × Female leader	−1	−3
	(6)	(5)
Constant	58***	166***
	(6)	(4)
Obs.	1569	1560
Clusters	39	39
Wald χ^2 / F statistic	2	1

B. Statistical analysis

Table A2 Table A2. Effort and earnings in Part 1 depending on the leader-selection procedure and the leader's gender in period 9contains regressions testing whether there are treatment differences in Part 1. In column I, the dependent variable is the participants' effort choice. It presents marginal effects (in percentages) of the probability of choosing an effort of zero in a given period based on an order probit regression. In column II, the dependent variable equals the participants' earnings in a given period. It presents coefficients from an OLS regression. In both regressions, we use as independent variables the interaction of treatment (*Election* or *Random*), the gender of the leader selected in period 9. By testing for differences in the treatment, we can corroborate whether random assignment indeed resulted in no differences in team performance. By testing for differences in the gender of the leader, we test whether the experience of teams in Part 1 determines what type of leader the group ends up having. We test for treatment differences using Wald test based on the estimated marginal effects. In both regressions, we cluster standard errors on teams. Note that the omitted category is teams who had a male leader in period 9 in *Random*.

Table A3 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

Table A3. 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	27	12	16	19	32	30
	s.d.	(18)	(17)	(19)	(16)	(19)	(20)	(16)	(17)
% coordinating on the highest effort	mean	30	20	67	17	37	46	78	75
	s.d.	(47)	(41)	(48)	(38)	(48)	(50)	(41)	(44)
Individual effort	mean	21	15	28	18	19	23	34	32
	s.d.	(19)	(19)	(18)	(18)	(19)	(19)	(13)	(16)
Earnings	mean	184	173	219	183	198	202	221	221
	s.d.	(72)	(77)	(44)	(68)	(57)	(59)	(53)	(48)
% expecting all will follow the leader	mean	16	20	19	13	18	22	29	28
	s.d.	(36)	(40)	(39)	(34)	(39)	(42)	(45)	(45)
% preferring a female leader	mean			32	72			23	74
	s.d.			(47)	(45)			(42)	(44)

depending on the treatment, the gender of the leader, and whether the current leadership term is the first term with a leader or not. Means and standard deviations are calculated using team outcomes per period for the teams' minimum effort and coordination on the highest effort. For individual effort and earnings, means and standard deviations are calculated using individual outcomes per period.

Table A4 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 include as control variables a dummy variable indicating the leader's race (Caucasian or non-Caucasian), a variable measuring the mean effort in Part 1 of the individuals' team, and dummy variables identifying

Table A4. 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	−52***	−37**	−25
	(18)	(15)	(20)
Random	−39*	−37**	2
	(21)	(17)	(19)
Random × Female leader	−50**	−49**	−7
	(22)	(21)	(25)
Periods 12 to 26	12	3	23*
	(10)	(8)	(13)
Periods 12 to 26 × Female leader	5	−2	18
	(14)	(9)	(13)
Periods 12 to 26 × *Random*	−32*	−21*	1
	(19)	(12)	(16)
Periods 12 to 26 × *Random* × Female leader	−28	−21*	2
	(20)	(13)	(17)
Caucasian leader	−8	−3	−6
	(9)	(7)	(8)
Team minimum effort in Part 1	3	3	1
	(3)	(2)	(2)
Constant	68***	218***	71***
	(15)	(10)	(15)
Obs.	702	3510	234
Clusters	39	39	39
R^2 / Pseudo R^2	16	7	11

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 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*. Next, we present robustness checks for some of the results seen in Table A4. In particular, we test whether the gender differences observed in the first leadership term are robust to alternative specifications. We ran these robustness checks because choices are highly correlated across periods following a leader's message.

Table A4-R. Team coordination, earnings, and individual effort depending on the leader-selection procedure and the leader's gender in the first leadership term

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

	I	II	III	IV
Female leader	−50**	−39**	−51***	−34***
	(19)	(16)	(18)	(9)
Random	−37*	−40**	-37*	−10
	(22)	(19)	(20)	(10)
Random × Female leader	−47*	−52**	−47**	−13
	(24)	(21)	(23)	(12)
Caucasian leader	−3	−20	−4	13*
	(16)	(14)	(15)	(8)
Team minimum effort in Part 1	−1	2	−1	−4**
	(3)	(2)	(3)	(1)
Won the lottery				0
				(10)
Constant	69***	228***	70***	70***
	(19)	(14)	(18)	(8)
Obs.	39	39	39	195
Clusters	39	39	39	195
R^2 / Pseudo R^2	19	24	15	9

In the first two columns, Table A4-R presents coefficients from OLS regressions and standard errors clustered on teams. In column I, the dependent variable equals the mean number of times a team coordinates on the highest effort in periods 9 to 11. In column II, the dependent variable equals the mean earnings in periods 9 to 11. In the last two columns, the table presents regressions utilizing data solely from period 9. Since we use data only from the first period after participants observe the leaders' message, in these regressions we do not cluster standard errors on teams. In column III, the dependent variable equals one if a team coordinates on the highest effort and zero otherwise. In column IV, the dependent variable equals one if an *individual* chooses the highest effort and zero otherwise. These two columns present marginal effects (in percentages) from probit regressions. In all regressions, we use as independent variables the interaction of treatment (*Election* or *Random*) and the gender of the leader in a period (male or female). In addition, we include as control variables a dummy variable indicating the leader's race (Caucasian or non-Caucasian) and a variable measuring the mean effort in Part 1 of the

Table A5. 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
Random × Female leader	−1	−7
	(28)	(21)
Election × Male leader	64***	−39***
	(18)	(13)
Election × Female leader	5	6
	(26)	(20)
Caucasian leader	−32*	−21
	(19)	(20)
Team minimum effort in Part 1	−7	−1
	(14)	(10)
Won the lottery		0
		(10)
Constant	36**	61***
	(18)	(13)
Obs.	117	516
Uncensored obs.	69	276
Clusters	39	39
Wald χ^2	389	387

individuals' team. In column IV, we also control for whether a participant received a random wealth shock because he or she won the lottery in the volunteering stage. Note that the omitted category is teams with a male leader in *Election*.

We can see that findings corresponding to the initial gender differences are robust to these alternative specifications. Specifically, in *Random* we do not find a significant difference between male and female leaders: $p = 0.345$ for mean coordination in the highest effort in periods 9 to 11; $p = 0.303$ for mean earnings in periods 9 to 11; $p = 0.339$ for coordination in the highest effort in period 9; $p = 0.403$ for choosing the highest effort in period 9. By contrast, we find that male leaders are significantly more effective than female leaders in *Election*: $p = 0.006$ for mean coordination in the highest effort in periods 9 to 11; $p = 0.010$ for mean earnings in periods 9 to 11; $p = 0.008$ for coordination in the highest effort in period 9; $p < 0.001$ for choosing the highest effort in period 9.

Table A6. Expected team coordination 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 the followers' expected probability of team coordination on 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
Prefers female leader	0	–8
	(6)	(16)
Female leader	–26*	
	(15)	
Prefers male leader × Female leader		–35*
		(20)
Prefers female leader × Female leader		6
		(20)
Caucasian leader	–34***	–32**
	(8)	(14)
Team minimum effort in Part 1	3	5
	(10)	(3)
Constant	73***	76***
	(10)	(5)
Obs.	96	96
Uncensored obs.	52	52
Clusters	24	24
Wald χ^2	5	10

Table A5 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 gender of the leader (male or female). In both regressions, we also include as control variables a dummy variable indicating the participants' race (Caucasian or non-Caucasian) and a variable measuring the mean effort in Part 1 of the individuals' team (in both the first and second stage of the regression), we cluster standard errors on teams, and restrict the regressions to data from periods 9 to 12. In column II, we also control for whether a participant received a random wealth shock because he or she won the lottery in the volunteering stage. Note that the omitted category is teams with a male leader in *Random*.

Table A6 presents marginal effects (in percentages) from probit models with sample selection testing the effect of a leader's gender and the participants' preference for the leader's gender on the followers' expectation of team coordination in *Election*. The regressions correct for the observation that follower expectations react positively to the leader's request of the highest effort, and leader requests can be affected by their gender and preference for the gender of a leader. In the regressions, for a given period, the dependent variable equals one if a follower expects that all other followers in the team will follow the leader's suggestion to choose the highest effort and zero otherwise. The first stage consists of a probit regression with a dependent variable that equals one if a leader asks for the highest effort in a given period. We use as independent variables, in both the first and second stage of the regression, the gender of the leader (male or female) and the participants' preference for a male or a female leader, which is measured by looking at the gender of their highest-ranked candidate in the preceding election (excluding themselves). Note that in the first election, all teams have at least three candidates. In both the first and second stage of the regression, we also include as control variables a dummy variable indicating the participants' race (Caucasian or non-Caucasian) and a variable measuring the mean effort in Part 1 of the individuals' team, we cluster standard errors on teams, and restrict the regressions to data from period 9. In column I, we present the estimated marginal effects of the gender of the leader and the followers' preference for the leader's gender but without considering the interaction between them. In column II, we include the interaction as an independent variable in both stages of the regression. Note that the omitted category is follower in teams with a male leader who have a preference for a male leader.

Table A7 presents marginal effects (in percentages) 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 columns I and II, the dependent variable equals one if a follower decides to become a candidate in the current election and zero otherwise. In columns III and IV, the dependent variable equals one if a candidate wins the current election and zero otherwise. In all regressions, we use, as independent variables, dummy variables to identify the gender of the

Table A7. Probability that followers challenge and that the leader is reelected depending on the leader's gender and team coordination on the highest effort

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

	I	II	III	IV
Unsuccessful female leader	6	4	−42***	−39***
	(5)	(5)	(15)	(9)
Successful male leader	−21***	−26***	32**	33***
	(5)	(4)	(13)	(9)
Successful female leader	−26***	−21**	19	13
	(9)	(8)	(17)	(19)
Caucasian leader	−2	−3	5	5
	(5)	(7)	(3)	(5)
Team minimum effort in Part 1	1	–	1	–
	(1)		(1)	
Constant	49***	–	54***	–
	(4)		(13)	
Obs.	482	482	93	93
Clusters	24	24	23	23
Wald χ^2	32	28	46	24

leader in the previous three periods interacted with whether the team was successful or unsuccessful (i.e., whether the team coordinated on the highest effort). Note that the omitted category is a previous male leader of an unsuccessful team. In all regressions, we also include as control variables a dummy variable indicating the race of the leader in the previous three periods (Caucasian or non-Caucasian) and a variable measuring the mean effort in Part 1 of the individuals' team. In columns I and III, the marginal effects are based on probit models, while in columns II and IV, they are based on conditional logit models with team fixed effects. In all 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). 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).

From the table, we can see that the followers of successful leaders are significantly less likely to volunteer than those of unsuccessful leaders for both male (48% vs 28%, $p < 0.001$ in both columns I and II) and female (56% vs. 26%, $p = 0.003$ in column I and $p = 0.030$ in column II) leaders. A difference-in-difference test reveals that the effect of team success on the probability

Table A8. Probability of being a leader depending on the participant's gender

Note: Marginal effects in percentages from a probit regression with being selected to be a leader 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	IV
Change if female	−14	2	−7	−11**
	(10)	(7)	(5)	(6)
Mean probability if male	25***	17***	21***	23***
	(6)	(4)	(3)	(4)

that a follower volunteers is not significantly different across genders ($p = 0.367$ in column I and $p = 0.722$ in column II). Similarly, successful leaders have a significantly higher chance of being reelected than unsuccessful leaders for both men (88% vs 58%, $p = 0.022$ in column I and $p = 0.113$ in column II) and women (78% vs. 14%, p<0.001 in column I and in column II). In this case, a difference-in-difference test reveals that the effect of team success on the probability that a follower volunteers is significantly stronger for male than for female leaders ($p = 0.029$ in column I and $p = 0.071$ in column II).

Table A8 presents marginal effects (in percentages) from a probit regression testing whether the probability of being a leader depends on one's gender. The dependent variable of the regression equals one in a period if an individual is the leader of his or her team and zero otherwise. The independent variables consist of dummy variables identifying the individual's gender (male of female), the treatment (Random or Election), and whether it is the first selection period (period 9) or a subsequent selection period (periods 12, 15, 18, 21, and 24). We include these variables with all interaction terms. In addition, we also control for the individual's race (Caucasian or non-Caucasian) and the teams' mean minimum effort in Part 1. Standard errors are clustered on teams. To facilitate the presentation of the marginal effects, we separate them into four columns. In columns I and II, we present marginal effects of the first selection period in either *Random*, in column I, or *Election*, in column II. In columns III and IV, we present marginal effects of the subsequent selection periods in either *Random*, in column III, or *Election*, in column IV. For each column, we present the estimated mean probability of being a leader if the individual is male and the estimated change in this probability if the individual is female instead. The regression uses 1170 observations, standard errors are clustered on the 39 teams, and the value of the χ^2 test of joint significant equals 17. The estimated marginal effect of the participant's race is small and is not statistically significant (3%, $p = 0.444$), while the marginal effect of the team's mean minimum effort in Part 1 is significantly positive ($p = 0.026$), which implies that teams who coordinated more in Part 1 were less likely to have a situation in Part 2 where nobody volunteers to be a leader.

We find that in the first selection period (period 9), there are no significant gender differences, suggesting that initially men are not more likely to be leaders then women. In the subsequent selection periods, a gender difference emerges. In *Election*, the probability that an individual is the leader of the team is 25% for men and 13% for women, the latter being significantly lower ($p = 0.041$). The same pattern is observed in *Random* but the gender difference is not statistically significant ($p = 0.122$). However, if we do a difference-in-difference comparison by testing whether the coefficients for female are significantly different between *Random* and *Election*, we find that they are not ($p = 0.547$). Hence, we cannot fully discard the possibility that the fraction of women leaders decreases only in *Election*.

C. Additional findings

Here, we present additional findings that are not reported in the paper due to space constraints.

We start by looking at the potential effect of the fraction of men and women in a team. In particular, research on leadership sometimes reports that the gender makeup of groups can affect the behavior of leaders and followers (e.g., Grossman, Komai, and Jensen 2015). In our experiment, participants were randomly assigned to teams and therefore, there is variation in the number of men and women in each team. Specifically, we have 5 teams with four men and one women, 13 teams with three men and two women, 19 teams with two men and three women, and 2 teams with one man and four women. For our analysis, we then distinguish between the 18 teams that have a majority of men and the 21 teams that have a majority of women. Importantly, the gender composition of the teams does not correlate the treatment (χ^2 test $p = 0.561$).

Table A9 presents marginal effects (in percentages) from probit regressions testing the effect of the teams' gender composition on team coordination, individual effort, the likelihood that a leader requests the highest effort, and the likelihood that the leader is male or female. In column I, the dependent variable equals one if a team coordinates on the highest effort in the given period and zero otherwise. In column II, the dependent variable equals one if an individual chooses the highest effort in the given period and zero otherwise. In column III, the dependent variable equals one if a leader asks the followers for the highest effort in the given period and zero otherwise. In

Table A9. Team coordination, individual effort, leaders' message choice, and leaders' gender depending on the leader-selection procedure and the gender composition in the team

Note: Marginal effects in percentages from a probit regression with the following dependent variables: (I) team coordination on the highest effort, (II) individual choice of the highest effort, (III) leader requests the highest effort, and (IV) an individual is selected to be the leader. Standard errors corrected for clustering on teams in parenthesis. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

	I	II	III	IV
Random × female	18	14	8	−4
	(24)	(22)	(18)	(10)
Election × male	43*	38*	19	2
	(23)	(20)	(16)	(4)
Election × female	28	26	13	−6
	(26)	(24)	(20)	(6)
Random × male × female majority	6	4	4	7
	(26)	(23)	(18)	(5)
Random × female × female majority	−10	−9	−8	−2
	(29)	(26)	(17)	(8)
Election × male × female majority	6	7	9	6
	(13)	(11)	(7)	(6)
Election × female × female majority	1	−2	−7	0
	(21)	(20)	(16)	(6)
Constant	32	41**	67***	21***
	(20)	(18)	(15)	(3)
Obs.	702	2823	234	1170
Clusters	39	39	39	39
Wald χ^2	11	13	13	14

column IV, the dependent variable equals one if the individual is selected to be the team leader in the given selection period and zero otherwise. In all regressions, we use as independent variables the three-way interaction of treatment (*Election* or *Random*), gender (male or female) of the leader in columns I, II and III and of the individual participant in column IV, and whether the team has three or more women or not (female majority or male majority). In addition, we use dummy variables to identify periods in which teams had no leader (which occurred in less than 3 percent of all periods). 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 regressions in column III and IV include data from every third period in periods 9 to 26.

We find that the team's gender composition does not have a statistically significant effect on any of the dependent variables. This can be seen from the marginal effects of the bottom four interaction terms, which are all relatively small and do not reach statistical significance.

Table A10. Volunteering to be a leader over time depending on the leader-selection procedure and the leader's gender

Note: Marginal effects in percentages from a probit regression with volunteering to be a leader as the dependent variable. In all cases, we subtract nine from the period number because there is the first period volunteering is possible. Standard errors corrected for clustering on teams in parenthesis. ***, **, and * indicate statistical significance at 1, 5, and 10 percent.

	I	II	III
Election	3		
	(5)		
Volunteering period	−3***	−3**	−6***
	(1)	(1)	(1)
Election × Volunteering period	−1***		
	(0)		
Female		−14	2
		(9)	(8)
Female × Volunteering period		0	3
		(2)	(2)
Constant	65***	73***	69***
	(04)	(5)	(4)
Obs.	1170	450	720
Clusters	39	15	24
Wald χ^2	99	19	81

Next, we look in more detail at the participants volunteering decision. Table A10 presents marginal effects (in percentages) from probit models testing whether the rate of volunteering changes over time. In all regressions, the dependent variable, in each period, equals one if an individual volunteers to be a leader and zero otherwise. In column I, we use as independent variables the interaction of treatment (*Election* or *Random*) and the number of volunteering periods (from 1 in period 9 to 6 in period 24). In column II and III, we use as independent variables the interaction of gender (male or female) and the number of volunteering periods, and we restrict the regression to either *Random*, in column II, or *Election*, in column III. In all regressions, we cluster standard errors on teams.

We find that in both treatments, volunteering significantly decreases over time but it decreases significantly faster in *Election* compared to *Random*. This is consistent with the paper's findings that successful leaders in *Election* are often not challenged by the followers in the subsequent election period. In columns II and III, we see that the interaction between gender and volunteering period is not significant, suggesting that the decrease in volunteering does not depend on the participants' gender.

Table A11. Expected payoffs from volunteering to be a leader depending on the volunteering decision of others

Note: Expected payoffs are calculated assuming that participants are rational own-payoff maximizers, are risk neutral, and are considering only one volunteering decision that will impact team play for the next three periods.

	Volunteer	**Do not volunteer**
0 others volunteer	$x + 50 + \alpha$	$x + 25$
1 other volunteers	$x + 25 + \alpha$	$x + 25 + \alpha$
2 others volunteer	$x + 16.7 + \alpha$	$x + 25 + \alpha$
3 others volunteer	$x + 12.5 + \alpha$	$x + 25 + \alpha$
4 others volunteer	$x + 10 + \alpha$	$x + 25 + \alpha$

Despite the decreasing trend, it is notable that we observe so many participants choosing to volunteer to become leaders. In the first leader selection period (period 9), 72% of participants volunteer in *Random* and 78% in *Election*. By the last leader selection period, volunteering rates have decreased to 51% in *Random* and 39% in *Election*. If we divide participants depending on whether they were the leader in the previous three periods, we find that 46% of previous followers volunteer in *Random* and 28% in *Election*, while 71% of previous leaders volunteer in *Random* and 87% in *Election*.

To shed some light into the participants' volunteering behavior, we calculate as a benchmark the Nash equilibria of the volunteering decision in the simplest of settings. Specifically, we assume that participants are rational own-payoff maximizers, are risk neutral, and are considering only one volunteering decision that will impact team play for the next three periods. We also assume that if there is no leader, then team members will earn x points over the next three periods, and that having a leader benefits the team by a total of α points over the next three periods. We start with the simple case, the *Random* treatment, where the probability of becoming the leader is the same for all who volunteer. Under these assumptions, Table A11 summarizes the payoffs of each participant faces depending on the number of other team members who volunteer. As we can see, if nobody volunteers then each team member can expect to receive x points plus 25 points, which is the expected prize of the lottery participants play if they do not volunteer (50 points with 0.5 probability and zero otherwise). As soon as one team member volunteers, then everyone receives the benefits of having a leader and those who volunteer receive an expected prize of 50 points divided by the number of volunteers.

This game has various Nash equilibria. To describe them it is useful to define a set V, which contains the team members who, in equilibrium, choose to volunteer with a strictly positive

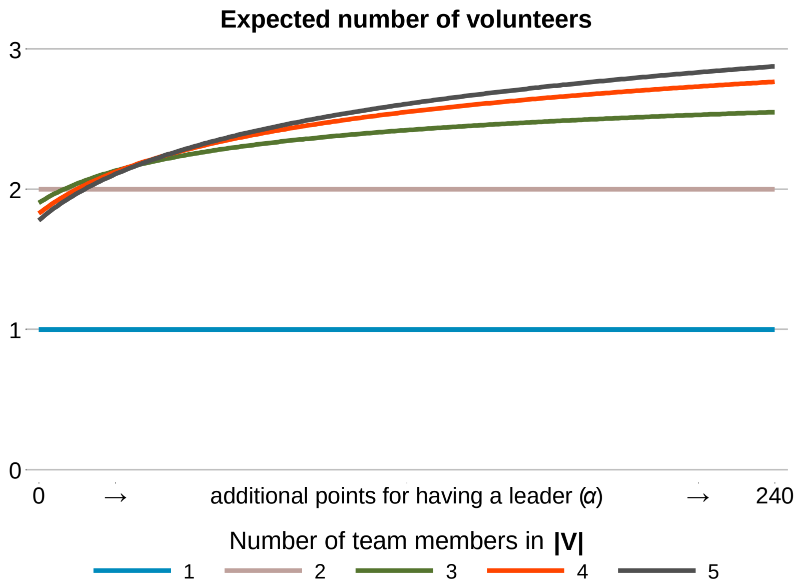


Figure A2. Average number of candidates in equilibrium depending on the number of team members in the set of team members who volunteer with positive probability $|V|$

probability. The average number of volunteers that occur in equilibrium depending on the benefits of having a leader, $\alpha \in [0,240]$, and the size of V are depicted in Figure A2. For $|V| = 1$ and $|V| = 2$, equilibria occur in pure strategies irrespective of the value of α.[1] In equilibria with $|V| = 1$, one team member volunteers with certainty and receives a higher payoff than others, who are indifferent between volunteering and not volunteering. In equilibria with $|V| = 2$, two team members volunteer and all team members receive the same expected payoff, but it is the volunteering team members who are indifferent between volunteering or not while non-volunteering team members have a dominant strategy to not volunteer. For $|V| \geq 3$, there are no longer any equilibria in pure strategies. There are, however, equilibria in mixed strategies where team members in V volunteer with a probability $p^*(a) \in (0,1)$, where $\partial p/\partial a > 0$. Note that the most efficient equilibria are those with $|V| = 1$, since they guarantee a leader and maximize the number of prizes obtained through not volunteering.

Comparing the predictions of this simple model to the participants' initial volunteering behavior, we can see they volunteer too much. This is very clear in the first period in which leaders are introduced. In period 9, the average number of volunteers per team is 3.60 participants in *Random* and 3.92 in *Election*. If we take the average difference in earnings between

[1] We are describing only cases where leaders have a positive effect on payoffs $\alpha \geq 0$. For $25 < \alpha < 0$, there are incentives to become a leader, even though it hurts the team to do so, as long as nobody else volunteers. For $\alpha < -25$, nobody volunteering is the Pareto-optimal equilibrium.

periods without a leader and the first three periods with a leader (49.65 points in *Random* and 106.65 in *Election*) as the value of α, the model predicts at most an average of 2.30 volunteers per team in *Random* and 2.56 in *Election* (predictions for $|V| = 5$).

Over time, the average behavior of participants falls within the model's predictions for two reasons. First, the value of a leader increases. For example, the average difference in earnings between periods without a leader and the last three periods with a leader is 119.67 points in *Random* and 176.67 in *Election*. For these values of α, the model predicts an average number of volunteers between 1.00 and 2.61 in *Random* and between 1.00 and 2.75 in *Election* (predictions for $|V| = 1$ and $|V| = 5$ respectively). Second, the observed number of volunteers decreases. By the last leader selection period, we find an average of 2.53 volunters per team in *Random* and 1.96 in *Election*. Hence, it appears that participants need some time to learn what the value of having a leader is and the volunteering behavior of their team members.

The extremely high initial volunteering rate is harder to explain. We think that is unlikely that it is the result of risk aversion since the amount of variation in payoffs for those who volunteer and those who do not is small. Moreover, given that initial volunteering rates are similar in *Random* and *Election*, it is hard to argue that high volunteering rate in Election are driven by overly optimistic beliefs of being elected. This leaves the possibility that participants initially think that they will be more effective leaders than their other team members. If this is a widely held belief, it is easy to imagine how it can result in the observed volunteering behavior.

References

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

Grossman, Philip J, Mana Komai, and James E Jensen. 2015. "Leadership and Gender in Groups: An Experiment." *Canadian Journal of Economics/Revue Canadienne d'économique* 48 (1): 368–88.

Reuben, Ernesto, and Krisztina Timko. 2018. "On the Effectiveness of Elected Male and Female Leaders and Team Coordination." *Journal of the Economic Science Association*.