The Emergence of Male Leadership in Competitive Environments

Ernesto Reuben[a], Pedro Rey-Biel[b], Paola Sapienza[c],
and Luigi Zingales[d]

[a] Columbia University, USA

[b] Universitat Autònoma de Barcelona, Spain

[c] Northwestern University, USA

[d] University of Chicago, USA

ABSTRACT

We present evidence from an experiment in which groups select a leader to compete against the leaders of other groups in a real-effort task that they have all performed in the past. We find that women are selected much less often as leaders than is suggested by their individual past performance. We study three potential explanations for the underrepresentation of women, namely, gender differences in overconfidence concerning past performance, in the willingness to exaggerate past performance to the group, and in the reaction to monetary incentives. We find that men's overconfidence is the driving force behind the observed prevalence of male representation.

1. Introduction

An important and often forgotten aspect of organizations is how efficient they are in selecting the right individual to perform a particular task. In competitive environments, an organization that is good at identifying and selecting its most talented individuals will be able to outperform organizations that might very well have more talented people but for some reason do not select them to perform the tasks they are good at. In this paper, we study experimentally whether organizations fail to select high-ability women to perform in competitive environments. Specifically, we investigate three potential gender differences that can lead groups to fail to select their high-ability women: first, gender differences in judging one's own abilities (e.g., due to overconfidence); second, gender differences in the willingness to exaggerate one's own abilities, and third, gender differences in how individuals respond to conflicts of interest between their own interests and their group's (i.e., agency problems). To do so, we ran an experiment where groups select a leader to represent their group and compete against the leaders of other groups in a real-effort task. Everyone in the group of the winning leader receives a prize. Barring any explicit discrimination against women—which would be unlikely in an experiment with university students—groups should aim to select their most talented individual irrespective of gender. However, we find that women are selected to represent the group 33.3% less often than their abilities would suggest. In other words, we find underrepresentation of women in a competitive environment that is not due to overt discrimination or to gender differences in preferences.

Our experimental design allows us to isolate the potential effect of each of the three aforementioned gender differences in order to explain why groups fail to select female leaders. The experiment was carried out with subjects who had already performed the real-effort task 15 months before participating in the current experiment. We took advantage of this fact to ask subjects, using an incentive compatible procedure, to recollect their past performance. If there is a gender difference in the recollection of past performance and groups are unaware of it, then low-ability but overconfident men will end up being chosen over high-ability but underconfident women.[1] In addition to remembering their past performance, we asked subjects to make known

---

[1] Throughout the paper, we follow the literature on self-serving attribution and define "overconfidence" as the overestimation of outcomes related to one's abilities (see Malmendier and Tate, 2008). This is also referred to as

to their groups what their expected performance is, although their statements did not have to be truthful. For groups to make the best decision, individuals must truthfully reveal their true ability to others. Therefore, if there is a gender difference in the willingness of individuals to exaggerate their own ability (e.g., compared to women, men might dislike revealing they have a low ability, or they might be more willing to lie and overstate their own abilities) then groups could end up selecting low-ability men that claim to be high-ability over high-ability women that do not overstate (or even understate) their ability. Finally, we introduce an agency problem by giving an additional financial incentive to the individuals who are selected as leaders. If, compared to women, men react more to the monetary incentive (and are willing to lie about their ability), then we ought to see less women as chosen to represent their group in treatments where the additional payment is large.

We find that the underrepresentation of women is mainly driven by men being overconfident when they recall their own ability. Once subjects of both genders have recalled their performance, they do not overstate their performance differently in order to be picked as the group leader. Finally, although higher economic incentives do induce subjects to overstate their performance, they have a similar effect on both men and women.

There is a growing literature on gender differences in competitive environments. Gneezy, Niederle, and Rustichini (2003) and Gneezy and Rustichini (2004) find that the relative performance of women vis-à-vis men worsens when they are competing against another individual, and in particular if the other individual is a man. Niederle and Versterlund (2007) demonstrate that women can be absent from competitive environments because they `dislike' being in competitive settings (compared to men). That is, women select themselves out of competitive environments even though they would be financially better off if they decided to compete. Dreber, von Essen, Ranehill (2009), Gneezy, Leonard, and List (2009) and Günther et al. (2010) demonstrate that the negative effects of competition on performance and selection might be the result of cultural differences and stereotypes associated with the particular task used. Niederle, Segal, and Vesterlund (2009) and Balafoutas and Sutter (2010) explore the effects of various policy interventions in lieu of gender differences in competitive environments. A survey of this literature is found in Booth (2009). This paper extends this line of research to

---

"optimism". Thus, for the purpose of this paper, the two terms can be considered to be interchangeable. Note that we do not define "overconfidence" as the underestimation of confidence intervals, as is done by some authors.

the case where selection into competitive environments is the outcome of a group decision. Finally, our paper is related to the work of Babcock and Laschever (2003) and Bowles, Babcock, and Lai (2007), who demonstrate that women tend to negotiate less than men and are more likely to be punished for negotiating. However, unlike in negotiation games, which tend to be zero-sum, in our setting incentives are more aligned and thus bias against selecting qualified women hurts not only these women but also the whole group.

The paper is organized as follows. In Section 2, we describe the experiment's design and procedures. In Section 3, we present the experiment's results. We provide a short discussion and conclusion in Section 4.

2. Experimental design and procedures

The experiment reported in this paper was run in the spring of 2008 with subjects from the Chicago-Templeton MBA Longitudinal Study (CTMLS).[2] A total of 134 subjects (100 males and 34 females) participated in the experiment. Subjects were randomly allocated to three sessions, which were run simultaneously in different classrooms within the University of Chicago.

For the experiment, we utilize the fact that all the subjects had previously participated in a first study conducted 15 months earlier in the fall of 2006. The 2006 study was part of a required class taken by all MBA students in the entering cohort and consisted of one survey and several games. For the purpose of this paper, we make use of the fact that in one of those games, subjects had to complete an adding task.[3] Specifically, subjects had 150 seconds to add as many sets of four two-digit numbers as they wished. They were not allowed to use a calculator, but could use paper and pencil to make calculations. The four (integer) numbers were randomly drawn from a uniform distribution with a support of 11 to 99. Once an answer was submitted, a new set of numbers appeared on the screen along with a message indicating whether their previous answer had been correct or not. The subjects' payment depended on the

---

[2] The CTMLS is a long-term research project on individual characteristics and economic success. It utilizes data from the 2008 cohort of the University of Chicago Booth School of Business. A detailed description of the CTMLS is available in Reuben, Sapienza, and Zingales (2008).

[3] The game, which is referred to as the "competition game" in Reuben, Sapienza, and Zingales (2008) is a replication of the experiment of Niederle and Versterlund (2007).

number of correct answers.[4] At the end of this study, all subjects were informed of their performance in the adding task. A detailed description of this study is available in Reuben, Sapienza, and Zingales (2008).

The current experiment consisted of two parts, one of which was chosen at random at the end of the experiment to determine the subjects' earnings. In the first part, subjects were reminded of the adding task they performed in the 2006 study and were asked to recall the number of additions they answered correctly. Specifically, we asked them for their performance during the piece-rate payment scheme. The subjects' payment in this part of the experiment depended on the accuracy of their answer: if their answer was within one addition of their actual performance, they earned $50 and otherwise $0. After reading the instructions for the first part, subjects submitted their answers individually before reading the instructions for the second part. The instructions for both parts are available in the Appendix.

For the second part of the experiment, subjects were divided into 33 groups of two to four. Groups were formed with stratified randomization by gender in order to ensure that we had enough groups with at least one woman.[5] In each group, subjects had to choose one group member to be their *representative*. They knew that the group's representative would compete against four other representatives by performing the same adding task as in the 2006 study. The representative with the highest number of correct answers would win the competition (ties were broken randomly), which gives every member of the winning group $4 per correct answer. The members of the losing groups earned $0. Regardless of the competition's outcome, representatives received a flat fee, which varied between groups: it equaled $0, $20, or $75.[6]

Subjects had five minutes of face-to-face communication to agree on who would be the group's representative.[7] The only restriction imposed on the subjects' communication was that

---

[4] As in Niederle and Versterlund (2007), subjects were paid either a small amount per question with certainty or a large amount per question if they happened to be the best performer in their four-person group.

[5] Due to time constraints and the fact that the experiment was not computerized, the randomization was done before the subjects arrived and was based on the subjects that signed up to participate. Groups that have less than four members are due to subjects not showing up to the experiment. Also because of absences, some of the groups consisted of only men. Excluding these groups from our analysis does not change our results.

[6] Of the 33 groups, 12 groups had a $0 fee, 9 groups had the $20 fee, and 12 groups the $75 fee.

[7] Subjects were told that groups that did not reach an agreement within the allotted time would not compete and therefore earn $0. Nonetheless, all groups managed to nominate a representative.

before they could talk freely, each group member had to state and write on a recording sheet their expected performance in the addition task. Note that subjects could state any number they wished and their answer in the first part of the experiment remained private information. The fact the subjects' were asked to remember their piece-rate performance but record their performance in a competitive environment gives subjects some wiggle room to think their performance will be different this time around (e.g., due to optimism). Once the five minutes were over, subjects wrote down their group's final agreement in the same recording sheet. Subsequently, representatives competed with each other in an adjacent room.[8] Earnings were calculated and subjects received their payment one week later during a mandatory class.

3. Results

For our analysis, we have 120 subjects divided into 33 independent groups.[9] Of these 33 groups, 29 of them (87.9%) chose a male representative and 4 of them (12.1%) chose a female representative. Given that women comprise 24.2% of the subjects, it appears that they are chosen as representatives half as often as their frequency in the population would suggest. A likelihood-ratio χ^2 test confirms that the likelihood of a women being a representative is significantly different than the likelihood of a men being a representative ($p = 0.046$).

Surely, there can be numerous reasons why men have a higher likelihood of being a representative. Assuming that groups make optimal decisions implies that the underrepresentation of women is solely the result of women being typically outperformed by a man in their group. Alternatively, gender differences in recollection of their past performance or in the willingness to claim a high performance can distort the group's decision and result in a suboptimal fraction of female representatives. To separate the effect of these explanations on the underrepresentation of women, we calculate the fraction of women that would be chosen as representatives in three hypothetical scenarios. In the first scenario, we assume that individuals know the *past* performance of all their group members and base their selection exclusively on who had the highest past performance. In the second scenario, we assume that individuals know the *remembered* performance of their group members and base their choice

---

[8] The addition task was programed with z-Tree (Fischbacher, 2007).

[9] Four groups did not follow the instructions and failed to state their individual claims before starting to talk, and therefore, we excluded them from the analysis. All these four groups chose a male representative.
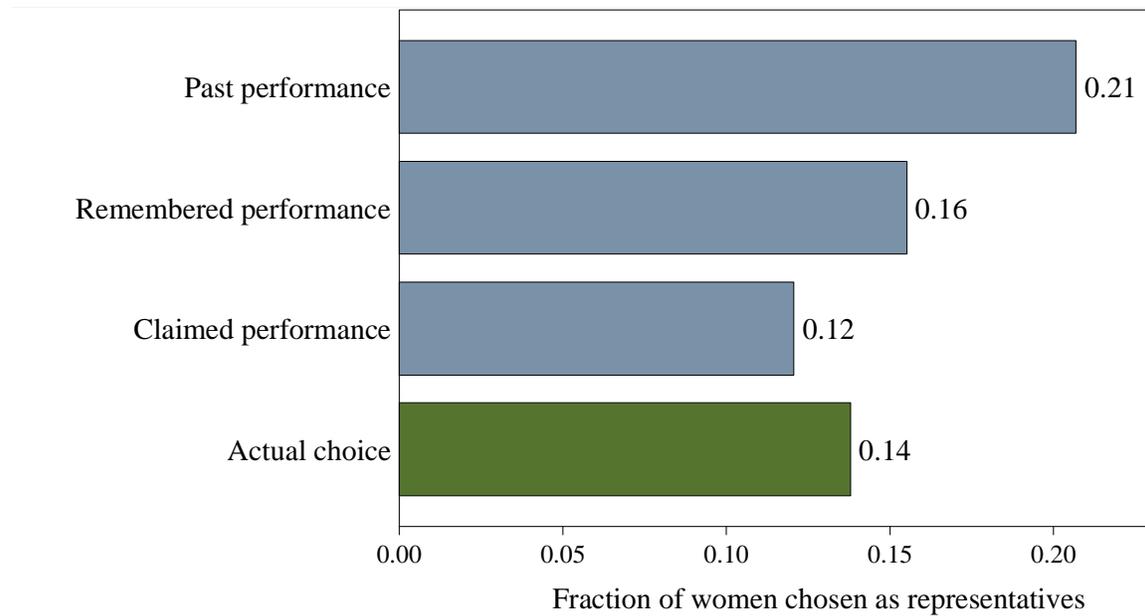
Past performance — 0.21

Remembered performance — 0.16

Claimed performance — 0.12

Actual choice — 0.14

Fraction of women chosen as representatives

Figure 1 – Fraction of women that would be chosen as representatives for different criteria

Note: The figure shows the fraction of women that were chosen as representatives and the fraction that would have been chosen if groups based their decision solely on the individuals': past performance, remembered performance, and claimed performance.

exclusively on it. In other words, the second scenario eliminates the distortion introduced by the self-reporting of performance but not the one induced by inaccurate recollection.[10] Lastly, in the third scenario, we assume that groups base their decision solely on the claimed performance of group members. That is, no new information or considerations are introduced during the five minute discussion. In all scenarios, we assume ties are broken randomly. The results of this exercise are presented in Figure 1.

As we can see in the figure, differences in past performance do not explain the small fraction of women chosen as representatives. If groups knew and based their decision on its members past performance, a woman would have a 20.7% probability of being a representative, compared to 13.8%, which is the actual fraction of female representatives. This probability falls considerably in the second scenario, where we dropped the assumption that individuals perfectly remember their past performance. Based on remembered performances, women have

---

[10] AS mentioned before, we refer to biases in remembering your own performance as overconfidence (or underconfidence). However, since the term "overconfidence" is usually used when individuals are forecasting the future and not when they are recalling the past, some readers will prefer the term "optimism".

Table 1 – Probability of being chosen as a representative

Note: Probit regressions with a dummy variable indicating whether a subject was chosen as a representative as the dependent variable. The table reports marginal effects and, in parenthesis, robust standard errors clustered on groups. In all regressions dummy variables are used to control for group size (not reported). All regressions have 120 subjects and 33 groups. * and ** indicate statistical significance at the 5% and 1% level.

	(I)	(II)	(III)	(IV)
Female	−0.203*	−0.223**	−0.106	0.019
	(0.077)	(0.051)	(0.127)	(0.112)
Best past performance		0.653**		
		(0.010)		
Best remembered performance			0.748**	
			(0.081)	
Best claimed performance				0.737**
				(0.070)
Log pseudolikelihood	−67.259	−41.841	−32.946	−32.844

a 15.5% probability of being a representative. In the third scenario, where representatives are chosen based on their claimed performance, this probability falls further to 12.1%, which is not far from the groups' observed choices.

To determine whether women face a statistically significant disadvantage in each of the three scenarios, we ran a series of probit regressions. As the dependent variable, we use a dummy variable indicating whether a subject was chosen as a representative (=1) or not (=0). As independent variables, we use the subject's gender, and dummy variables indicating whether the subject was the best performer in the group according to each of the three scenarios (past performance, remembered performance, and claimed performance). Finally we use dummy variables to control for the effect of group size. The results are presented in Table 1.

As expected, we can see that without controlling for performance, women are at a disadvantage: they have a significantly lower probability of being a representative (column I). In column II we see that, although having the best past performance in a group significantly increases an individual's probability of being chosen, it hardly affects the women's lower probability of being a representative. In contrast, once we control for whom has the highest remembered performance, the coefficient for female, although negative, is no longer statistically significant (column III). Finally, as one would anticipate from looking at Figure 1,
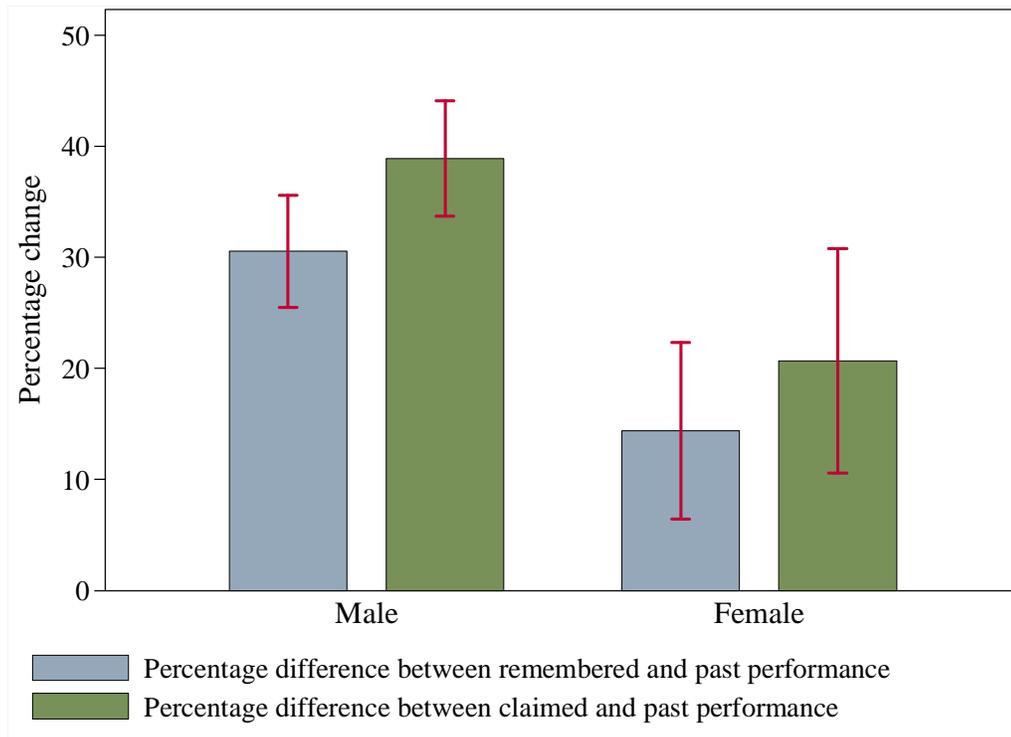
Figure 2 – Overestimating and overstating their past performance

Note: The figure shows the mean percentage difference between the subjects' actual past performance and the performance they remembered and claimed 15 months later. Error bars correspond to ± one standard error.

women are no longer discriminated against once we control for who claims to have the highest performance (column IV). This last observation is revealing in the sense that it indicates that groups are not explicitly discriminating against women (i.e., they never chose a man who claimed a lower performance than a woman). Instead, women are not being chosen because, from the groups' point of view, women have lower performances. Judging by the regressions in Table 1, this seems to be driven by differences in the way men and women remember their past performance. We explore this conjecture in further detail next.

On average, when estimating their past performance, subjects overestimated the number of additions they had answered correctly. Male subjects remembered answering 2.345 additions more than they actually did and female subjects remembered answering 0.917 additions more than they actually did. As seen in Figure 2, these statistics correspond to male subjects overestimating their performance by 30.5% and female subjects by 14.4%. Running an OLS regression with the difference between past and remembered performance as the dependent variable and a gender dummy as the independent variable confirms that, although

8

both men and women significantly overestimate their past performance ($p < 0.001$ for men and $p = 0.070$ for women), men do so to a larger extent ($p = 0.026$).[11] Interestingly, the proportion of subjects overestimating their performance is roughly the same for both men and women (70.5% vs. 64.3%, a likelihood-ratio χ^2 test gives $p = 0.542$). In other words, compared to women, men are not more likely to overestimate their own performance, but conditional on doing so, they overestimate it by a much larger margin.[12] Next, we look at the subjects claimed performance.

When making their statement to their group concerning their own performance, subjects claimed they will do better than their remembered performance. On average, subjects claimed they will answer 0.473 more additions than they remembered doing. Around half of the subjects (50.9%) claimed precisely the number of additions they remembered, about one third (31.3%) claimed a higher number of additions, and around one fifth (17.9%) a smaller number of additions. As one would expect, the discrepancy is driven by groups in which representatives received an additional fixed fee. If the representative's fee was $75, subjects claimed they will answer 0.927 more additions than they remembered doing, if the fee was $20 they claimed 0.364 more additions, and if there was no fee, they claimed only 0.079 more additions. This result is consistent with Gneezy (2005) who shows that individuals lie more when the positive economic consequences of their deception are larger.[13]

Unlike for remembered performance, we do not find a gender difference if we look at the difference between the number of claimed and remembered additions. On average, when representatives did not receive a fee, men claimed 0.100 more additions and women claimed 0.000 more additions. This suggests that men are not more prone than women to exaggerate their own performance when there are no financial incentives to do so. In groups where representatives did receive a fee, both men and women increase the distance between their

---

[11] The estimated regression equation is: remembered – past performance = 2.345 (0.384) – 1.428 (0.631) × female. Robust standard errors are in parenthesis.

[12] If instead of remembering their exact past performance subjects are simply stating focal-point guesses, then it is possible that women have lower estimates because they use a lower focal point (e.g., 10 vs. 15 additions). We do not see evidence that this is the case: the modal choice is 7 additions for women and 8 additions for men, neither of which are very focal, and we do not see spikes in the estimates distribution at focal numbers such as 5, 10, or 15.

[13] In fact, given that subjects are part of the same MBA program and thus might know each other, the effect of the fee is bound to be a lower bound as subjects ought to be more reluctant (and have a smaller incentive due to potential sidepayments) to lie to their friends.

claimed performance and remembered performance: men claimed 0.667 more additions and women claimed 0.706 more additions. However, we do not find a gender difference in how men and women react to the financial incentive.[14] Lastly, we find that similar proportions of men and women claimed *fewer* additions than the ones they remembered (overall, 19.5% for men and 12.0% for women). If one interprets this behavior as a consequence of subjects not wanting to represent their group in the subsequent tournament, then we do not see that women are purposefully avoiding being a representative and thus from competition in the experiment. The fact that we do not see a gender difference in competitiveness, unlike Niederle and Versterlund (2007) and subsequent replications, could be the result of: subjects not realizing that they could strategically manipulate their answer to opt out of the tournament, a self-selection of highly competitive women in the MBA program, or the possibility that preferences for competition are affected by whether one is competing for oneself or as part of a group (see Healy and Pate, 2011).

In summary, as seen in Figure 2, an analysis of the subjects remembered and claimed performance confirms that the main difference between men and women is that men overestimate their past performance by a larger extent, which explains the higher frequency of male representatives. There are multiple reasons why we see this gap in expectations. On the one hand, it could be due to men being more overconfident or optimistic than women such that they recall performing better than they actually did. On the other hand, it might be driven by individuals (correctly) believing that women will perform worse than men in a competitive environment, in particular if they compete against men (see Gneezy, Niederle, and Rustichini, 2003; Gneezy and Rustichini, 2004).[15] Given that too few women are selected as representatives, we cannot test whether their performance suffers when they compete against men, and thus, we cannot exclude that there is such an effect. We can, however, check whether representatives were in fact too optimistic with respect to their subsequent performance.

---

[14] These results are confirmed with an OLS regression with the difference between claimed and remembered performance as the dependent variable, and the following independent variables: dummy variables to control for gender and the size of the representatives' fee. The coefficient for the gender dummy variable is not statistically significant ($p = 0.989$), and compared to no fee, the coefficient for the $75 fee is significant at the 10% level ($p = 0.061$). Moreover, there is no significant interaction between gender and the representative's fee.

[15] Given that in our experiment men outnumber women by a significant margin, the belief that most other competitors would be men is well justified.

To conclude with this section, we take a look at the representatives' actual performance in the inter-group competition. After all, we have shown that groups deviate from the choice they would have made if they based their decision solely on the individuals' past performance. However, it is possible that the subjects' remembered performance and/or claimed performance are actually better variables for groups to base their decision on. We find that this is not the case. The correlation coefficient between the representatives' actual performance and their claimed performance is 0.406 and the correlation coefficient with their actual and remembered performance is 0.506. Both coefficients are significantly lower than the correlation coefficient between their actual and past performance, which is 0.864 ($p < 0.001$ with Hotelling-Williams tests). We also see that the subjects' overestimation of their past performance does not translate into a higher present performance. Specifically, on average representatives answered 10.879 additions, which is very similar to the 10.687 additions they answered 15 months earlier.[16] However, they remembered answering significantly more: 14.424 additions ($p < 0.001$ with a paired t test). These results confirm that subjects are indeed overconfident, and thus, that overconfidence is a likely explanation for the underrepresentation of women as representatives. Accordingly, if a group (but no others) is able to overcome the biases observed in the data, it would be able to significantly improve the selection of its representative and consequently their performance in the inter-group competition.[17]

4. Conclusions

In this paper, we report experimental evidence showing that groups select women to perform in competitive environments much less often than their abilities would suggest. We explore three potential explanations for this underrepresentation of women. We find evidence that suggests that gender differences in overconfidence concerning their own performance explain a significant proportion of the lack of female leadership. In contrast, we do not find support for other potential explanations: namely, gender differences in the wiliness to exaggerate past performance or in the reaction to monetary incentives.

---

[16] This means we do not find a difference in performance between the piece-rate and tournament payment schemes.

[17] A rough back-of-the-envelope calculation indicates that a group that overcomes these biases will gain on average around 0.500 additions in the inter-group competition.

Our findings are important as they highlight mechanisms that can lead not only to underrepresentation of women, but also to decisions that are suboptimal for the group. They also point at the type of policy interventions that could or could not improve the situation. For example, notice that in our context, overconfidence refers to a bias in individuals' recollection of their past performance, which they use to form an assessment of their future performance. In other words, it is a bias that can affect group decisions even when everyone's incentives are perfectly aligned. In order to overcome this problem, individuals would have to be given reliable measures of their own past performance. Along these lines, we can also expect to see that this type of bias is more pronounced in populations that are prone to overconfidence, as perhaps MBA students are, compared to other populations. We leave these questions to future research.

Our results are consistent with literature that studies the effect of on overconfidence on the decisions of individuals in leadership roles. Malmendier and Tate (2005; 2008) show that CEOs—a group where women are underrepresented (Bertrand and Hallock, 2001; Wolfers, 2006)—not only possess high degrees of overconfidence but also that this distorts their investment decisions. Our experiment suggest that this can be the outcome of group decision-making that does not fully take into account differences in overconfidence across groups and individuals.

References

Babcock, L., and Laschever, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide.* Princeton: Princeton University Press.

Balafoutas, L. and Sutter, M. (2010). Gender, Competition and the Efficiency of Policy Intervention. Working Papers in Economics #450. Göteborg University.

Bertrand, M., and Hallock, K. (2001). The Gender Gap in Top Corporate Jobs. *Industrial and Labor Relations Review,* 55(October): 3-21.

Booth, A. (2009). Gender and competition. *Labour Economics*, 16(6): 599-606.

Bowles, H.R., Babcock, L., and Lai, L. (2007). Social Incentives for Gender Differences in the Propensity to Initiate Negotiations: Sometimes It Does Hurt to Ask. *Organizational Behavior and Human Decision Processes*, 103(1): 84-103.

Dreber, A., von Essen, E., and Ranehill, E. (2009). Outrunning the Gender Gap – Boys and Girls Compete Equally. Working Paper Series in Economics and Finance 709. Stockholm School of Economics.

Fischbacher, U. (2007). Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics,* 10(2): 171-178.

Gneezy, U. (2005). Deception: The Role of Consequences. *American Economic Review*, 95(1): 384-94.

Gneezy, U., and Rustichini, A. (2004). Gender and Competition at a Young Age. *American Economic Review*, 94(May): 377-381.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender differences. *Quarterly Journal of Economics,* 118(August): 1049-1074.

Gneezy, U., Leonard, K., and List, J. (2009). Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. *Econometrica,* 77(5): 1637-1664.

Günther, C., Ekinci, N.A., Schwieren, C., Strobel, M. (2010). Women can't jump?–An experiment on competitive attitudes and stereotype threat". *Journal of Economic Behavior & Organization,* 75(3): 395-401.

Healy, A. and Pate, J. (2011). Can Teams Help to Close the Gender Competition Gap? *The Economic Journal*, forthcoming.

Malmendier, U., and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance,* 60(6): 2661-2700.

Malmendier, U., Tate, G. (2008). Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction. *Journal of Financial Economics*, 89(1): 20-43.

Niederle, M., and Vesterlund, L. (2007). Do Women Shy away from Competition? Do Men Compete too Much? *Quarterly Journal of Economics*, 122(3): 1067-1101.

Niederle, M., Segal, C., and Vesterlund, L. (2009). How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." NBER Working Paper 13923.

Reuben, E., Sapienza, P., and Zingales, L. (2008). A Description of the Chicago-Templeton Longitudinal Study. Working Paper. University of Chicago.

Wolfers, J. (2006). Diagnosing Discrimination: Stock Returns and CEO Gender. *Journal of the European Economic Association,* 4(May): 531-41.

Appendix – Instructions

This section contains the instructions used in both parts of the experiment.

A.1 First part

In the fall of 2006, as part of LEAD, you participated in a series of games played on laptop computers. One of the games consisted of summing four two-digit numbers. A summary of the instructions of this game is seen below:

> ### Welcome to Game 3
>
> At the beginning of the game you will be divided into groups of four. You will be given a series of *addition tasks* (sums of four 2-digits numbers). You will have 150 seconds to answer as many questions as you want. The computer will record the number of sums that you answer correctly. You will be paid *$4* for each correct answer you give.

Next, you can also see a snapshot of the computer screen of the addition task.

Today's tasks

Today you will be asked to perform two tasks. One of the two tasks will be chosen at random to determine your payment for today.

Task 1

As mentioned above, in one case, you performed the addition task and where paid $4 for each correct answer. Please indicate below the number of sums that you remember getting correct in this instance. You will be paid according to the accuracy of your guess. Specifically, if your guess is within ± 1 of the correct number of answers you earn *$50*, else you earn $0.

My best guess of the number of correct answers is: _____

Indicate how confident you are of the accuracy of your guess:

completely unsure O O O O O O O completely certain

A.2 Second part

Task 2 – Your Group's Decision

For this task, you are assigned to a group of people. The task of the group is to select a representative to play the same game against the representatives of 4 other groups. The representative with the highest number of correct answers wins the competition.

Each member of your group will earn $4 per correct answer if your representative wins the competition. Otherwise everyone in your group earns $0. In addition each group's representative receives a flat fee of [...]. This fee is paid independently of whether your group wins the competition or not.

You will have 5 minutes to agree (with at least a 3 to 1 majority) on who will be the group's representative. If you do not agree, nobody will play and you will receive no money for the game.

When your group meets, each group member has to announce his/her expected performance, then each member of the group has to announce whether they are interested in being the representative and why. Thereafter, you can bargain freely. At the end you have to

write down the final agreement reached by the group (if any) and the votes supporting it. You need to record your decision in the RECORDING SHEET.

The representatives will perform the task during the break. Next week, you will be told the performance of your representative and of the representatives of the groups you are competing with.

RECORDING SHEET

Please fill in the following at the BEGINNING of the negotiation.

Group Member: [xxx]   Expected performance: _____
Reason to be / not be the group's candidate: _____

_____

Group Member: [xxx]   Expected performance: _____
Reason to be / not be the group's candidate: _____

_____

Group Member: [xxx]   Expected performance: _____
Reason to be / not be the group's candidate: _____

_____

Group Member: [xxx]   Expected performance: _____
Reason to be / not be the group's candidate: _____

_____

Please fill in the following at the END of the negotiation.
Final Agreement: _____

_____

_____

Representative chosen: _____
Names of the people who voted in favor: _____