Breaking the Glass Ceiling with "No": Gender Differences in Declining Requests for Non-Promotable Tasks[1]

Lise Vesterlund, University of Pittsburgh, Economics, and NBER
Linda Babcock, CMU, Heinz College and Department of Social and Decision Sciences
Laurie Weingart, CMU, Tepper School of Business[2]

Abstract: Gender differences in task allocations at work may contribute to the persistent vertical gender segregation that exists in the labor market. If women spend more time on non-promotable tasks (tasks that are not relevant for their advancement) then they may progress more slowly than men in organizations. Using laboratory experiments, we investigate in a stylized setting whether men more than women decline requests for non-promotable tasks. Our design captures the incentives members of a group face when asked to volunteer for a task that each member prefers that another member of the group undertakes (such as writing a report, serving on a committee, organizing an event etc.). We find that women volunteer twice as often as men. While the differential is not explained by individual characteristics such as risk and altruism, it does appear to be influenced by beliefs. Men and women are equally likely to volunteer when they know that they are grouped only with members of their own sex. This response to same-sex grouping shows that the individual's willingness to volunteer is not fixed, and it suggests that the documented gender gap in volunteering likely results from the belief that women are more likely than men to volunteer. Extending our study to the field we ask whether male and female faculty members differ in their willingness to serve on university committees. Data from a large public University replicate results from mixed-sex setting in the laboratory. Relative to men, women faculty are far more likely to accept the request to serve on a committee. We argue that differences in the propensity to say 'no' to requests for non-promotable tasks will result in different task allocations which in turn can create barriers to the advancement of women in organizations and in society as a whole.

[1] We thank Craig Fox and David Tannenbaum for giving us permission to use their data. We thank David Klinowski, Conor Lennon, Maria Recalde, and Amanda Weirup for superb research assistance. Participants at seminars and conferences at the University of Zurich, Stanford, UCSB, CMU, University of Pittsburgh and Stockholm School of Economics are thanked for their helpful comments. Finally we thank the Carnegie Bosch Institute and the NSF (SES-1330470) for generous financial support.

[2] In accordance with the norms in psychology, we order the authors' names according to their contributions.

1. Introduction

Despite significant female educational advances, we continue to see gender differences in labor market outcomes (Goldin, Katz and Kuziemko, 2006; Bertrand, Goldin, and Katz, 2010). Horizontal as well as vertical job segregation is substantial, causing men and women to have very different labor market experiences (Altonji and Blank, 1999; Bertrand and Hallock, 2001). Common explanations for these persistent differences are discrimination and that men and women differ in their abilities or preferences over jobs (e.g., Polachek, 1981; Goldin and Rouse, 2000; Black and Strahan, 2001). Another more recent explanation is that men and women differ in their response to the circumstances they face in the work place. In particular research documents that relative to men, women are more reluctant to negotiate (Babcock and Laschever, 2003) and to compete (e.g., Gneezy, Niederle and Rustichini, 2003; Niederle and Vesterlund, 2007). It has been argued that both of these differences contribute to the persistent vertical gender segregation.

This paper investigates whether differential responses to requests for non-promotable tasks may be another explanation for why women advance at a slower rate than men in the work place. In most professional occupations, workers have some discretion over the allocation of their time. A worker can allocate her time to tasks that are likely to improve her performance evaluation ("promotable tasks") or to tasks that, while benefitting the organization, are unlikely to affect her evaluation and advance her career ("non-promotable tasks"). As an example consider an untenured assistant professor at a research university. She knows that to be promoted, the best use of her time is to focus on her research. What should she do when asked by her Dean to be a member of the University Faculty Senate? She knows that this will take a lot of her time, take her away from her research, and be unlikely to produce any real rewards for her. Yet, important faculty matters are debated in the faculty senate and the institution is important to the well-functioning of the University. How will she respond to such requests? Will her response differ from that of a male colleague with comparable credentials?

Recent research indicates that women allocate more of their time to tasks that are less likely to lead to promotions. In a survey of 350 faculty at University of Amherst, Misra et. al (2011) find that compared to men, women faculty spent 7.5 fewer hours per week on research and 4.6 more hours per week on university service committees. Using a US sample of 1,399 political science faculty Mitchell and Heslie (2013) find that women advised more undergraduate students and participated in more department and college level committees. They conclude that women more than men provide "token" services that are unlikely to lead to advancement.

Exploring areas outside of academia, De Pater et. al (2010) find that among workers in mid-level jobs, men, more than women, evaluate their individual task assignments as challenging. Furthermore they show that the difference in part can be attributed to differential task assignments by a supervisor. A consequence may be that women spend more of their time doing more menial tasks that fail to provide them with the recognition needed to advance in the organization.

Gender differences in task allocations may result both from demand and supply.[3] On the demand side, men and women may differ in the frequency by which they receive requests to do non-promotable tasks. On the supply side, men and women may differ in the extent to which they offer to perform or accept requests to perform "non-promotable" tasks. In particular women may be more likely than men to agree to such requests if they are more other-regarding and concerned for the welfare of others (e.g., Eckel and Grossman, 1998; Andreoni and Vesterlund, 2001),[4] if they are more agreeable and have a greater desire to be liked by the requestor (Braiker, 2001), and if they have a greater desire to conform to a norm of accepting such requests (e.g., Santee and Jackson, 1982; Eagly, et al., 1981). Essential when deciding whether to decline a request is the assessment of the consequences of doing so. In particular women will be more likely than men to accept the request if they are more risk averse and concerned about the consequences from declining the request (e.g., Eckel and Grossman, 2008), or if they are less confident that others will undertake the undesirable, but necessary, task in their place.[5]

In trying to understand what may give rise to different task allocations we investigate if men and women differ in their 'supply' of non-promotable tasks. That is we examine whether men and women differ in their propensity to say 'no' to requests to perform tasks that they would prefer someone else does in their place.

As a starting point we conducted, along with Amanda Weirup, two pilot surveys to determine whether men and women differ in their perception of undesirable requests (Weirup et. al, 2013). Respondents were asked to describe work-related requests that they would have preferred to decline but nonetheless agreed to do. They were asked to report the factors that affected their decision to accept the request and the emotions they felt while making the decision. One survey used a small sample of MBA students (n=47) and asked them to describe one such request. A second survey used a larger sample of mTurk U.S. based participants (n=212) and asked participants to recall five work-related requests.[6] These data revealed that women were more likely than men to experience negative emotions (anger, fatigue, and guilt) and to be more worried than men about the negative repercussions of declining such requests. By contrast, men were more likely than women to experience positive emotions (happiness and calmness) and were more likely than women to consider instrumental reasons for agreeing to requests such as being owed future favors. As examples, women more frequently reported that they felt pressured to say 'yes,' that they found it very difficult to say no, and that they worried that someone would see them as not helpful. As an overall assessment of the propensity to say 'yes,'

[3] In discussing both the demand and supply of non-promotable work-related tasks we are considering tasks for which there is a request and some discretion over the acceptance of the request. Thus it may be argued that we examine tasks that lie between those considered in the psychology literature on "organizational citizenship behaviors" (OCBs) where individuals on their own initiate tasks that benefit the organization (e.g., Organ, 1988) and those considered in the organizational psychology literature's examination of task allocation where the employee must accept an assigned task (e.g. De Pater et al 2010).

[4] See also Eagly & Mladinic (1989), Eagly & Steffen (1984), and Heilman (2001) for evidence that women are more motivated than men to engage in behavior that helps others rather than themselves.

[5] See Croson and Gneezy (2009) for a careful review of the economics literature on gender differences.

[6] mTurk (Mechanical Turk) is a service platform provided by Amazon that allows requesters to post various tasks for workers to do on their computer for payment.

participants were asked to rate on a scale from '1' to '7' whether they agree with the statement that they 'agree to do too many things at work,' with '1' denoting strongly disagree and '7' strongly agree. Women were significantly more likely than men to agree with the statement (Mean for women 5.1 versus 4.3 for men, $p < .06$).

To secure promotable task assignments it is essential that a worker, when asked to do a task, can assess whether there is discretion to decline the task and whether it is in her interest to do so. An individual's work-place success depends on her ability to identify and decline tasks that are discretionary and non-promotable. Our survey suggests that relative to men, women are more concerned about the negative consequences of declining requests for such undesirable tasks and, that they find it more difficult to do so. The career consequences of accepting a discretionary and non-promotable assignment can extend beyond the assignment itself. In particular such tasks may generate lower job satisfaction and in turn can reduce the worker's commitment and investment in her job.[7]

Our survey finding, that men and women perceive requests differently, need however not imply that behavior differs as well. To determine whether the response to undesirable requests differs by gender, we conducted a laboratory experiment where participants in a group were presented with a task that only one person can undertake. The return from performing the task is such that the individual will only undertake it if no one else is willing to do it. Despite presenting men and women with identical incentives, we find that women are almost twice as likely to do the task. In investigating alternative explanations for this gender difference we find that the propensity to do the task is not fixed, but depends on the characteristics of the other potential volunteers. In particular the gender gap in the propensity to agree to the request is sensitive to whether the other members of the group are of the same or the opposite sex. Conducting the experiment using mixed or same-sex groups we document a response to gender pairing that is consistent with the belief that women are more likely than men to perform the task.

Finally to determine whether gender differences in declining requests also arise in the field, we examine data on volunteering to serve on University committees. These volunteering decisions reveal that female faculty are far more likely than male faculty to volunteer to serve on committees. Assuming that faculty prefer to have someone else serve on committees, these field data confirm our laboratory finding. In mixed-sex settings women more than men accept requests for tasks that, while helping the collective, are unlikely to advance the individual's career.

2. Experiment 1: Do men and women differ in their response to undesirable requests?

We designed a laboratory experiment to determine whether men and women differ in their response to requests for 'non-promotable' tasks. We wanted to capture the incentives a small group faces when it is asked to find a volunteer for a task that everyone is reluctant to undertake, be it write a report, serve on

[7] Finding that career interruptions and differences in weekly hours explain differences in salaries for male and female MBAs (Bertrand, Goldin, and Katz, 2010) need not imply gender inequality if differences in labor market attachment result from differential task allocations.

a subcommittee, plan a holiday party, etc. The setting we have in mind is one where every member of a committee or group prefers that the task be undertaken, yet everyone prefers that it be undertaken by someone other than themselves. With the request for a volunteer being made to the group, every member waits for a volunteer to step forward, fully aware that an excessive delay increases the likelihood that an inferior outcome will result (such as the task not being completed in time or not completed at all). As no explicit request is made of any one individual, the request is implicit and arises through time pressure.[8]

2.1. Design

Capturing the incentives described above, our experimental design is as follows. In each of ten rounds participants from a large pool are randomly and anonymously assigned to groups of three. Members of the group are then given 2 minutes to make an investment (volunteering) decision. Individual earnings are $1 in the event that no one invests before the end of the 2 minutes. If one group member makes the investment, the round ends, and the individual making the investment secures a payment of $1.25, while the other two group members each receive $2. The investor is randomly determined in the event that multiple parties simultaneously invest.[9]

With no cost of waiting, investments will be made in the last second of the round and the game reduces to one of simultaneous moves. Accounting for the possibility of ties, the game gives rise to three types of equilibria: Pure strategy asymmetric Nash equilibria where one individual invests and the others do not; a mixed strategy symmetric equilibrium where each player invests 23.2% of the time; and a mixed strategy asymmetric equilibrium where one person does not invest and the two others invest 40 percent of the time. Depending on the equilibrium selected the probability that an investment occurs is 100%, 54% or 64%, respectively.

2.2. Participants and Procedures

The experiment was conducted at the Pittsburgh Experimental Economics Laboratory (PEEL) at the University of Pittsburgh. Participants were recruited from introductory economics classes and were only informed that they would participate in a study on decision making. None of the participants had prior experience with studies at PEEL. The experiment lasted slightly less than an hour and average earnings were $22.8 (including a $6 show up fee).

Nine sessions were conducted, with between 12 and 21 participants per session, for a total of 150 participants (82 males and 68 females). Sessions were roughly gender balanced with the share of women participating in a session ranging between 33% and 53%.[10] The population was rather

[8] While delay does not carry a monetary cost it may carry a psychological cost. Bliss and Nalebuff (1984) examine a model with costly delay where individuals decide whether to secure the provision of a binary public good.

[9] The game may be characterized as a three person hawk-dove game where participants start off playing hawk and any participant's decision to play dove resolves the coordination problem.

[10] There is no evidence that this small variation in gender composition affects behavior. The likelihood that individuals invest is not affected by this degree of variation in gender composition. Clustering on the individual and controlling for round a probit of the individual's propensity to invest on the share of women in the session reveals

homogeneous. The average age was 18.9 years, with 18 and 19 year olds accounting for 76 percent of the participants. 74 percent were Caucasian, 87 percent were born in the US, and 85 percent were either freshmen or sophomores. None of these characteristics differed significantly by gender.[11]

Upon entering the lab, participants were seated in a pre-marked cubicle, and were asked to provide informed consent to participate in the study. We then distributed instructions and read them out loud. The instructions explained all procedures of the study, the payoff structure, the random matching protocol, and what information participants would receive during the study. Using a computerized interface (z-Tree, Fischbacher, 2007) we then began the ten round decision phase of the experiment.[12] In each round participants were anonymously matched in groups of three, with the stipulation that no one could be paired with the same person twice in a row. Each group member was shown an individual computer screen that displayed the seconds remaining in the round and a button that could be clicked if the individual wanted to invest. The round ended the second someone in the group clicked the investment button. Participants waited until all groups had either made an investment decision or the two minutes passed without an investment being made. Participants were again randomly matched into groups of three with the constraint that participants could not be matched with the same participant two rounds in a row.

At the end of the 10 rounds, participants answered a number of questions to assess individual preferences and characteristics. A demographic questionnaire elicited gender, age, nationality, year in college and college major. Gender was not mentioned until the very end of the experiment.

2.3. Results

To characterize behavior in the experiment we first ask whether groups succeeded in making investments and what the timing was of such investments. We then examine whether men and women were equally likely to make the investment. After documenting a substantial gender difference we determine whether it can be explained by differences in preferences for risk, altruism, conformity, and agreeableness.

2.3.1. Group Investments

Over the course of the ten rounds groups succeed in investing 84.2% of the time (std. error = .0183). With two thirds of these investments being made with less than 3 seconds remaining in the round, participants largely treated the environment as one of a war of attrition.

---

a marginal effect of 0.016 (p=0.959). Furthermore the marginal effects reported in our central results are not affected by the share of women in the session (the coefficient on female is 0.144 rather than the 0.141 seen in Table 1), nor is the coefficient on the session share of women significant (p=0.574).

[11] The mean age of men and women is not significantly different (18.98 vs. 18.78, p=0.319). Similarly using a Fisher's exact test there is no significant gender difference in the distribution of age (p=0.854), race (p=0.681), number of years in the US (p=0.587), years in college (p=0.292) or in choice of major (p=0.681).

[12] See Appendix A for the instructions and a sample decision screen.

Behavior in one session (session 7) differed substantially from the other eight sessions. Rather than delaying investments until the end of the round, participants in session 7 appeared to compete to invest first.[13] 35 percent of investments in session 7 were made in the first second of a round, by comparison this share did not surpass 6 percent in any one of the other sessions (session 1-6, 8, 9). Similarly no investment was ever made in the last second in session 7, whereas the percent of investments made with less than one second remaining ranged between 11 and 22 percent in the other eight sessions.[14] To assess the response to the intended war of attrition environment we opt to exclude session 7 from our analysis. In Appendix B we first demonstrate how behavior in session 7 differs from behavior in the other sessions and then show that our results are robust to the inclusion of these data.

Eliminating session 7 from the analysis, we are left with 8 sessions and a sample of 72 males and 60 females. With ten rounds and three people per group we examine a total of 440 group decisions. Common among these decisions is that investments are delayed until the end of a round, and that this delay becomes more prominent over the ten rounds. The rows in Table 1 report the number of seconds remaining at the time when an investment decision was made, the columns report how the distribution of decision times changes over the course of the experiment. Looking at the last two columns we see that the share of investments made in the last 2 seconds of a round increases from 63.3 percent during the first half of the experiment (Rounds 1-5) to 90.6 percent during the second half (Rounds 6-10). The bottom row of Table 1 shows the likelihood that a group succeeded in making an investment before the end of the two minutes. Looking across the columns we note that the likelihood that a group succeeds in investing decreases from 86.8% during the first half to 77.3% during the last half. The per round decrease in the investment rate is significant, starting at 93 percent in round one and ending with a 73 percent success rate in round 10.[15]

[13] In fact investments in session 7 were made so quickly that z-tree reported the time of investment as having 99999 seconds remaining of the round. This type of error result from the investment button being pressed before the clock can begin (see e.g. Duffy and Smith, 2011).

[14] Similarly in session 7 only 43 percent of investments were made in the last five seconds. Looking at each of the remaining eight sessions we note that in no session does the share of investments which were made in the last 5 seconds fall below 62 percent, with the share reaching 88 percent in one session.

[15] Treating the group as the unit of observation a probit regression of the probability that a group invests on round number yields a marginal effect of -0.02 (z=-5.94, p=0.001). See Figure 5 for the time trend in the likelihood that groups succeed in investing.

Table 1: Distribution of investment times

Seconds remaining at time of investment	Relative Frequency of investments		
	Round 1-10	Round 1-5	Round 6-10
Less than 1	14.4	5.2	24.7
1	49.0	37.2	62.4
2	12.7	20.9	3.5
3-10	5.8	9.9	1.2
11-20	0.3	0.5	0.0
21-30	0.3	0.0	0.6
31-40	0.8	1.6	0.0
41-50	0.8	1.6	0.0
51-60	0.8	1.0	0.6
61-70	0.8	1.6	0.0
71-80	0.3	0.5	0.0
81-90	0.8	1.0	0.6
91-100	1.9	3.1	0.6
101-110	3.0	4.2	1.8
111-120	8.0	11.5	4.1
Percent of groups investing	82.0	86.8	77.3
Total number of group decisions	440	220	220

2.3.2. Individual Investments: Gender difference

With four out of five groups securing an investment the group's rate of success exceeds that predicted by the symmetric mixed strategy equilibrium. This raises the question of whether certain members of the group more frequently make the mutually beneficial investment, and in particular whether the likelihood that the individual invests differs by gender. Figure 1 reveals that women are systematically more likely than men to undertake the investment.[16] Starting in round 1 the investment rate by women surpasses that by men. With a sustained differential over the ten rounds, this results in a substantial difference in the total number of times men and women invest over the ten rounds. Women on average invest in 3.5 of the 10 rounds, whereas men invest in 2.1 of the 10 rounds. This 67 percent difference in total investment is statistically significant ($t = 5.76$, $p<.001$). Figure 2 shows by gender the distribution of total investments over the ten rounds. The distribution for women first order stochastically dominates that for men, and the difference in the two is significant (a Fisher's exact test yields $p = 0.095$).[17] While 65 percent of men invest two or fewer units over the ten rounds, only 35 percent of women fall in this lower contribution range.

[16] As we only observe the individual making the investment we cannot compare the individual investment rate with the equilibrium prediction.

[17] The median contribution is 2 for men and 4 for women ($p = 0.002$).

Figure 1: Probability of investing

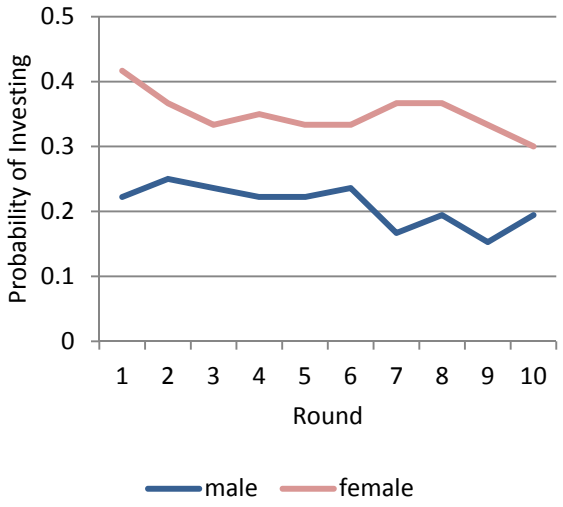
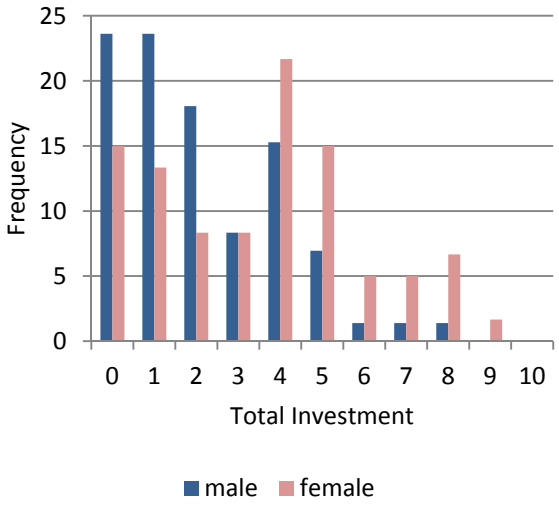


Figure 2: Distribution of total investment



In Table 2 below we estimate the probability that men and women invest in a given round. Standard errors are clustered on the individual to account for the lack of independence. The reported marginal effects confirm the insights from Figure 1 and 2. Pooling the data from all ten rounds we see in column 1 that participants become less likely to invest over the course of the experiment and that women are significantly more likely to invest than are men. The average investment rate for men is 21% and that for women is 14 percentage points higher. Columns 2 and 3 confirm that these results hold both for the first and second half of the experiment.[18] That is the gender differential is persistent over the course of the experiment.

Table 2: Probability of Investing (probit)[19]

	(1) All rounds	(2) Rounds 1-5	(3) Rounds 6-10
Female	0.141	0.129	0.151
	(0.000)	(0.001)	(0.002)
Round	-0.007	-0.010	-0.010
	(0.046)	(0.405)	(0.319)

Dependent variable: Individual investment decision (1-invest, 0-don't invest). The table presents marginal effects. Standard errors are clustered on the individual. p-values are reported in parenthesis. 132 participants

2.3.3. Individual Investments: The role of preferences

To understand the cause of the substantial gender gap in investment rates, we ask whether it may be accounted for by differences in preferences or attitudes. Conditional on the individual's belief that others invest, we expect participants to be more likely to invest if they are more risk averse, altruistic,

[18] The likelihood that a man invests is 21 percent over the ten rounds, 23 percent during the first five rounds, and 19 percent during the last five rounds.

[19] Including a (female x round) interaction results in a small but insignificant increase in the gender gap over the course of the experiment. The marginal effect of the interaction is 0.003 (p=0.660).

agreeable or conforming. We relied on the measures described below to elicit these individual characteristics.[20]

Risk Taking. We used both a survey and an incentivized measure to elicit the individual's willingness to take risk. For the survey measure participants used a 5-point scale to answer the Dohmen et al (2011) question: "How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks" (1 = not at all willing to take risks, 5 = very willing to take risks).[21] For the incentivized risk measure we used a version of the Eckel & Grossman (2002) measure: Pick your preferred 50-50 gamble from the following set of six (G1: $14, $14, G2: $12, $18; G3: $10, $22; G4: $8, $26; G5: $6, $30; G6: $1, $35). While both risk and the expected return increase when moving from the safe gamble G1 to gamble G5, only risk increases when moving from gamble G5 to G6.[22] As seen in earlier studies (for a survey see Croson and Gneezy, 2009) we find that men are significantly more risk taking than women (Men mean = 3.74, Women mean = 3.08, $p < .001$ for the survey measure; Men mean = 3.91, Women mean = 3.17, $p < .01$ for the incentivized measure).

Agreeableness. To elicit a measure of "agreeableness" we used the 9-item agreeableness subscale of the Big Five Personality Scale (John & Srivastava, 1995). For each item the participant assesses the extent to which each sentence applies to them, such as "I see myself as someone who is considerate and kind to almost everyone" and "I see myself as someone who starts quarrels with others" (reversed). Participants answered these questions on a 1 to 5 scale with 1 indicating strong disagreement and 5 indicating strong agreement. As is traditional in the literature we averaged the 9-items to form our agreeableness measure (α=. 0.8028). Consistent with previous research we find a higher agreeableness score for women than for men, the difference is however not significant (Men mean = 3.71, Women mean = 3.84, $p = 0.285$).[23]

Altruism. Our measure of altruism uses the three item principle of care construct by Wilhelm and Bekker (2010).[24] The principle of care construct aims to assess participants' feelings about helping others such as "People should be willing to help others who are less fortunate" and "These days people need to

[20] As an ex post measure of how participants perceived the experiment we also developed a 4-item scale that measured the extent to which participants had felt anxious during the decision phase. Example questions are: "How stressed did you feel during the decision task?" and "How anxious did you feel during the decision task?" Participants rated the extent to which they felt this not at all (1) versus extremely (5). There was a small and insignificant gender difference in this measure of anxiety (Men mean = 2.23, Women mean = 2.48, $p = 0.147$). We did not collect this measure for session 1 and 2. Inclusion of this ex post measure of anxiety does not affect any of our results.

[21] Dohmen et al (2011) develop this question and show that responses correlate with risky behaviors such as stock holdings, occupational choice, and smoking. In fact they show that it is a better predictor than an incentivized lottery. To keep the survey questions on a comparable scale we ultimately opted to have participants use a 5 point scale for all survey questions.

[22] Risk-neutral subjects are predicted to select G5 or G6, risk-seeking subjects should select G6. Those completely averse to risk should select the safe option G1, those less averse to risk should select G2, G3 or G4.

[23] See e.g., Schmitt et al (2008).

[24] Across a series of five studies Bekkers et al (2013) show that principle of care predicts charitable giving. Principle of care is included as a measure of altruism on the General Social Survey.

look after themselves and not overly worry about others" (reversed). Participants indicated their agreement with those items on a 1-5 scale with 1 indicating strong disagreement and 5 indicating strong agreement. The three items were averaged to form our altruism measure (α=0.594). We found no gender differences in the altruism scale (Men mean = 4.02, Women mean = 4.17, p > .24).[25]

Non-Conformity. Weber et al.'s (2002) social risk measure was used to elicit our measure of non-conformity. Including the three questions with the highest factor loadings, questions were of the form "How likely are you to admit that your tastes are different from those of your friends?" and "How likely are you to defend an unpopular issue that you believe in at a social occasion?" Participants indicated their agreement with those questions on a 1-5 scale with 1 indicating extremely unlikely and 5 indicating extremely likely. We averaged the scores on the three items to form our non-conformity scale (α=0.579). We found that men were significantly more non-conformist than women (Men mean = 3.85, Women mean = 3.43, p< .002).[26]

The correlations between our survey measures are shown in Table 3 below, and are broadly as expected. Our incentivized and survey measure of risk taking are positively correlated, and agreeableness is positively correlated with altruism and negatively correlated with non-conformity. Non-conformity and altruism are positively correlated.

Table 3: Pairwise correlation coefficients

	Survey risk	Lottery risk	Agreeable	Altruism	Nonconform
Survey risk	1				
Lottery risk	0.205*	1			
Agreeable	-0.091*	-0.049*	1		
Altruism	-0.106*	0.024	0.319*	1	
Nonconform	0.382*	0.108*	-0.185*	0.049*	1

*significant at the 5% level or better

While we find significant gender differences in some of our individual measures (risk-taking and non-conformity), these differences do not account for the observed gender gap in investments. In Table 4 we estimate the probability that men and women invest in a given round with and without controls for individual characteristics. Standard errors are clustered on the individual and the reported coefficients are marginal effects. Looking at column (2) we see that none of the elicited measures predict the individual's propensity to invest. Furthermore when controlling for individual characteristics the

[25] While a number of studies find that women are more altruistic, others find the opposite result. Andreoni and Vesterlund (2001) suggest that the gender gap in observed generosity depends on the cost of giving. The principle of care measure has previously been found to generate a small but insignificant gender difference. For example using a sample of 1163 households from the GSS Mesch et al (2011) find principle of care scores for men of 3.76 (0.61) and for women of 3.92 (0.63).

[26] Using a larger set of social risk questions Weber et al (2002) do not find a gender difference in social risk.

coefficient on female remains significant, and the comparison between columns (1) and (2) reveals that the magnitude largely is unchanged.[27]

Table 4: Probability of investing (probit)

	(1)	(2) [28]
Female	0.141	0.125
	(0.000)	(0.004)
Round	-0.007	-0.007
	(0.046)	(0.042)
Non-conform		-0.023
		(0.338)
Risk-seeking		-0.011
		(0.585)
Altruism		0.005
		(0.843)
Agreeable		-0.031
		(0.329)

Dependent variable: Individual investment decision (1-invest, 0-don't invest).
The table presents marginal effects. Standard errors are clustered on the individual.
p-values are reported in parenthesis.132 participants

In contrast to expectations we do not find that our measures of individual preferences help us explain the gender gap in the propensity to invest. One factor we have not investigated however is the role of beliefs. Do beliefs about whether men versus women will invest differ and can such a difference explain the gender difference in actual investing? To examine the role of beliefs and to determine if the response to requests is fixed we decided to conduct a second experiment where all participants in the session were of the same sex.

3. Experiment 2: Is the response to requests fixed?

If a woman believes that others expect her to invest and that they will not invest in her absence then it is in her best interest to invest, likewise if a man believes that the investment is secured without him then it is in his interest to abstain.

One potential way of investigating the role of beliefs is to manipulate them by changing the visible gender composition of the group.[29] If men and women are believed to differ in the likelihood by which

[27] Inclusion of other personal characteristics (age, race, years in school etc) does not affect the coefficient on female, and none of these factors are statistically significant.

[28] The risk measure included in column (2) refers to our survey measure of risk. Our incentivized risk measure was not elicited in sessions 1 and 2. In the reduced sample of 102 participants the incentivized risk measure is not significant, nor does it affect the coefficient on female. In the reduced sample the coefficient on female is 0.143 (p=0.005) when using the survey measure of risk and it is 0.148 (p=0.004) when using the incentivized risk measure.

[29] Examining coordination in the battle of the sexes game, Holm (2000) points to the role of a gender based focal point. He found that when males and females were paired with a female, they were more likely to select the action associated with their preferred equilibrium. This 'hawkish' behavior implied that coordination more frequently was achieved in mixed-sex pairing – thus securing higher average earnings. In our setting the

they invest, then the role of beliefs can be tested by examining whether the sex of the other group members influence the individual's decision to invest. We could investigate this by indicating the sex of the group members with whom a participant is matched. We were concerned that doing so would alert participants to our interest in how their decisions responded to the sex of their group members. We therefore opted to conduct a series of single-sex sessions. That is, instead of inviting close to equal proportions of men and women to our laboratory for each session, we instead conducted sessions where only men or only women participated. Our aim with these single-sex sessions was two-fold. First they will help to determine whether gender differences in investments are robust to the environment or if they are influenced by beliefs. Second it will allow us to determine whether, when recognizing that they are with members of their own sex, all-female groups are more likely than all-male groups to succeed.[30] Specifically if women, relative to men, are always more likely to agree to requests then the likelihood of group investments should be larger in the all-female sessions than in the all-male sessions. However, if the decision to invest is driven by beliefs about the behavior of others, then these beliefs may depend on the population from which group members are drawn. If women are expected to have a higher rate of investment, then men may increase the frequency by which they invest when they recognize that the other members of their group are men (that is, when they participate in the all-male session), and women may decrease the frequency by which they invest when they recognize that the other members of their group are women (that is, when they participate in the all-female session).

3.1. Participants and procedures

The recruitment method, instructions and procedures of these single-sex sessions were identical to those of the mixed-sex sessions in Experiment 1. 117 undergraduate students were recruited from introductory economics classes at the University of Pittsburgh and the characteristics of this pool of participants were similar to those of Experiment 1. Participants were on average 18.7 years old, with 18 and 19 year olds accounting for 81 percent. 74 percent were Caucasian, 91 percent were born in the US, and 85 percent were either freshmen or sophomores.

We conducted 3 sessions with all women (n=65) and 3 sessions with all men (n=52). Sessions consisted of between 15 and 24 participants. With ten rounds we have a total of 390 group decisions. Average individual earnings were $22 (including a $6 show-up fee).

Following the procedures of Experiment 1 we did not mention gender at any point during the instruction or decision making phase. Participants were only asked to report their gender on the survey at the end of the experiment. Although gender was not mentioned, by looking around the room participants could see the population from which members of their anonymously drawn groups would be drawn. Hence

investment can only be made by one individual. Hence coordination issues are resolved when investing, and the decision not to invest cannot be justified on the grounds that it reduces the chance of mis-coordination.

[30] In our first experiment participants did not know the gender composition of the group they were in in any particular round, however with women being more likely to invest all-female groups more frequently secured an investment than did groups that had at least one man in the group (All-female mean = 0.926, Not-all-female mean = 0.814, two-sided p = 0.141.)

the gender composition of the room had the potential of influencing the participants' beliefs about the likelihood that other members in their group would invest.

3.2. Results

Looking first at the aggregate results we pool the data from the all-male and all-female sessions. Aggregate behavior and investment times were similar to those of the mixed-sex sessions. Comparable to the 82% investment rate in the mixed-sex sessions of Experiment 1, we find an investment rate of 80.8% in the single-sex sessions of Experiment 2. As seen in Table 5 the timing of the investments is also similar. Investments were primarily made at the very end of the round. 75 percent of investments were made with less than 3 seconds remaining, and this share increased over the course of the experiment. In fact the share of investments made with less than one second remaining more than tripled between the first and second half of the experiment (increasing from 6.3 to 21.6 percent). Furthermore, over the course of the experiment the likelihood that a group invests decreases from 90.3 percent in the first half to 71.3 percent in the second half. A probit regression reveals that this decrease in investments over time is significant.[31]

Table 5: Distribution of investment times (Experiment 2)

Seconds remaining at time of investment	Frequency of investments		
	Round 1-10	Round 1-5	Round 6-10
Less than 1	13.0	6.3	21.6
1	43.8	40.9	47.5
2	18.4	22.2	13.7
3-10	5.4	8.5	1.4
11-20	0.0	0.0	0.0
21-30	1.3	0.6	0.0
31-40	1.3	1.1	2.2
41-50	0.6	1.1	1.4
51-60	1.0	1.7	0.0
61-70	1.0	1.1	0.7
71-80	0.3	0.6	0.0
81-90	0.6	1.1	0.0
91-100	1.3	1.7	0.7
101-110	1.9	2.3	1.4
111-120	10.2	10.8	9.4
Percent of groups investing	80.8	90.3	71.3
Total number of group decisions	390	195	195

3.2.1 Individual investments: Gender difference

The likelihood and timing of investments in Experiment 2 is very similar to that in Experiment 1. Next we ask whether these aggregate results mask differences between the all-female and all-male groups. Of particular interest is whether the gender gap in the probability of investing is sensitive to the single-sex setting.

[31] Treating the group as the unit of observation a probit regression of the likelihood that a group invests on round number yields a marginal effect of -0.038 (p=0.001). See Figure 5 for the time trend in the likelihood that groups succeed in investing.

Figure 3 displays by gender the probability that an individual invests in a given round. In sharp contrast to our results from Experiment 1 we do not find that women are any more likely to invest than men. The average number of investments over the ten rounds does not differ by gender (Men mean = 2.67, Women mean = 2.7, t = .17, p>.86).[32] When the gender composition of the group is known to the participants there is no evidence that all-female groups fare better than all-male groups. The success rate is 81 percent for women and 80 percent for men.

Figure 3: Probability of investing (Exp. 2)

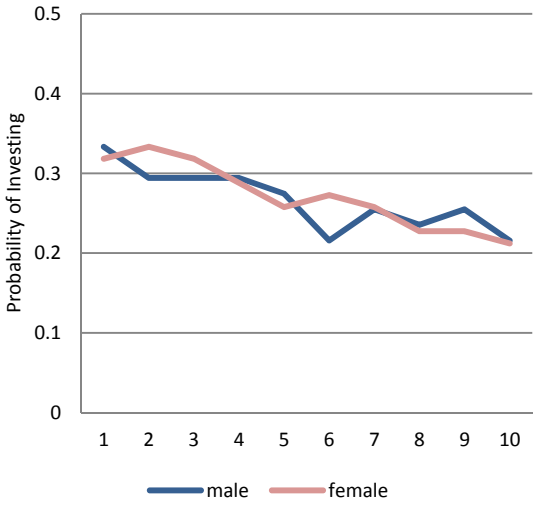


Figure 4: Distribution of total investment (Exp. 2)

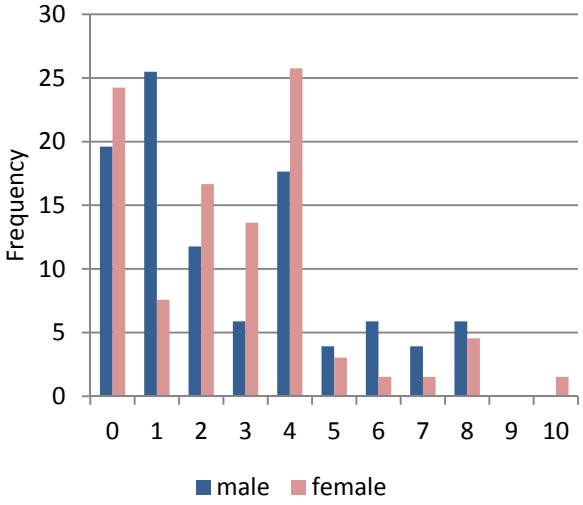


Figure 4 shows the distributions of total investments over the ten rounds by gender. It is interesting to note that although the aggregate investment rates are independent of gender, the two groups secure these investment rates in two very different ways. More than half of the women select a 'fair' investment rate and invest two to four times over the ten rounds. Men on the other hand are more likely to not invest (zero or one time) or to invest very often (five or more times over the ten rounds).[33] Despite the similarity in aggregate investment rates, the burden of these investments are more unequally distributed among the men that it is among the women.[34]

Verifying the results of Figure 3 we see from the probit models in Table 6 that in single-sex sessions the decision to invest does not depend on the sex of the participant. Across all ten rounds, the first five rounds, and the last five rounds, we find that men and women are equally likely to invest. The coefficient on Female is small in magnitude and imprecisely estimated.

[32] The median contribution is 2 for men and 3 for women (p = 0.368).
[33] While the absence of women does not affect the share of men who rarely invests (0-1 times) it causes some men to become very generous.
[34] A Fisher's exact test for equality of distributions yields a p = 0.158.

Table 6: Probability of Investing (probit)

	All rounds	Round 1-5	6-10
Female	0.004	0.005	0.004
	(0.920)	(0.916)	(0.935)
Round	-0.012	-0.015	-0.009
	(0.000)	(0.251)	(0.391)

Dependent variable: Individual investment decision (1-invest, 0-don't invest). The table presents marginal effects. Standard errors are clustered on the individual. p-value is reported in parenthesis. 117 participants.

The finding that there is no gender difference in the probability of investing when participants make decisions in single-sex groups helps explain why differences in individual characteristics did not explain the differential seen in the mixed-sex treatments. The changes in investment rates between Experiment 1 and 2 demonstrate that the individual's behavior is not fixed, but rather a function of the population from which group members are drawn. This suggests that beliefs about the behavior of others play a central role when deciding whether to undertake the undesirable task. Interestingly individual characteristics may play a larger role when there are less tangible attributes to coordinate on, such as an individual's reluctance to undertake risk. As seen in Table 7 we find that a measure such as risk aversion helps predict investments in the single-sex setting.

Table 7: Probability of Investing (probit)[35]

	(1)	(2)
Female	0.004	-0.002
	(0.920)	(0.962)
Round	-0.012	-0.013
	(0.000)	(0.000)
Non-conform		0.038
		(0.188)
Risk-seeking		-0.050
		(0.044)
Altruism		0.035
		(0.295)
Agreeable		0.020
		(0.614)

Dependent variable: Individual investment decision (1-invest, 0-don't invest).
The table presents marginal effects. Standard errors are clustered on the individual.
p-value is reported in parenthesis. 117 participants

[35] To be consistent with Table 4 we rely on the survey measure of risk. Using instead the incentivized measure of risk the coefficient on female would be -0.014 (p=0.762) and on risk -0.027 (p=0.134).

3.2.2 Individual Investments: Sensitivity to gender pairing

Figure 5 summarizes the results of the two sex-composition treatments. The group's rate of success is independent of the group being drawn from a single- or mixed-sex population, and in the single-sex sessions it is independent of whether the group is all-male or all-female.[36]

As demonstrated in Figure 6 the individual's propensity to invest is sensitive to treatment. Single_m and single_f refer to the all-male and all-female sessions (Experiment 2), respectively. Mixed_m and mixed_f refer to males and females in the mixed-sex sessions (Experiment 1), respectively. Relative to the mixed-sex results the investment rate decreases for women and increases for men in single-sex sessions.[37]

Figure 5: Probability group invests

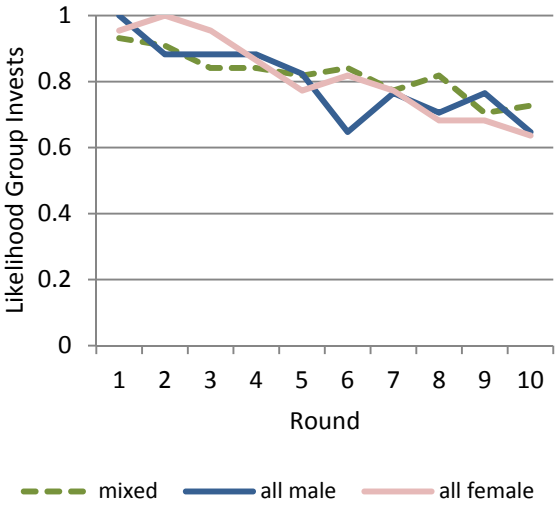
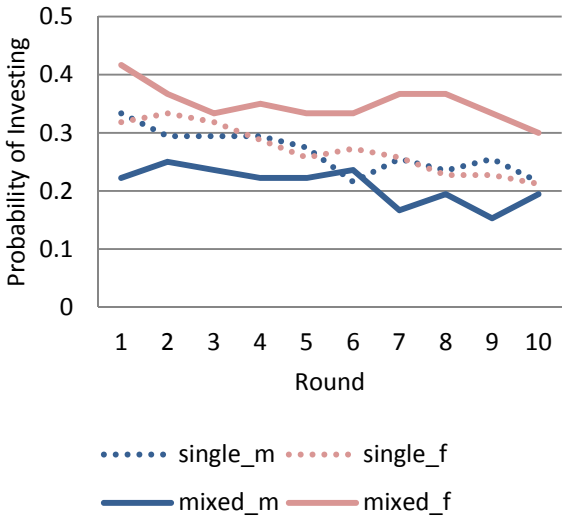


Figure 6: Individual propensity to invest



The differential response of men and women in the mixed-sex session is consistent with women being pessimistic and men being optimistic that other members of their group contribute. The single-sex results demonstrate that the frequency of investments changes when participants realize their group consists solely of members of their own sex. To summarize we examine in Table 8 the effect of gender composition on the gender gap in investments. While women are significantly more likely to invest in

[36] Treating the group decision as the unit of observation we see for single-sex session data alone that controlling for round a probit of group investment on female generates a marginal effect of 0.016 (p=0.735). Similarly pooling all the data and including a dummy for single-sex sessions we get a marginal effect of -0.010 (p=0.716).

[37] The average total investment by women is 3.5 in the mixed-sex session and 2.7 in single-sex sessions (p = 0.067). While one may be concerned that this differential results from 3.5 surpassing the level achievable in a single-sex session we note that the standard error on the 2.7 average total investment is 0.28. The average total investment by men is 2.1 in the mixed-sex session and 2.7 in single-sex sessions (p = 0.147). Separating the sample by gender a probit regression of the decision to invest on treatment (1=single-sex, 0=mixed-sex) and round, yields a marginal effect on single-sex of -0.08 (p=0.064), for the subsample of men we see instead an increase in the likelihood that an investment is made with the marginal effect on single_sex being 0.057(p= 0.153). Standard errors in both regressions are clustered on the individual.

mixed-sex sessions, this gender gap is eliminated in single-sex sessions, and this differential arises both in the first and the second half of the experiment.[38]

Table 8: Probability of Investing (probit)[39]

	All rounds	Round 1-5	Round 6-10
Female	0.139	0.129	0.147
	(0.000)	(0.001)	(0.002)
Round	-0.010	-0.012	-0.009
	(0.000)	(0.162)	(0.188)
Single-sex	0.061	0.071	0.051
	(0.152)	(0.117)	(0.299)
FemalexSingle	-0.136	-0.124	-0.147
	(0.021)	(0.041)	(0.034)

Dependent variable: Individual investment decision (1-invest, 0-don't invest). The table presents marginal effects. Standard errors are clustered on the individual. p-value is reported in parenthesis. The coefficient and standard error on the interaction is corrected to account for the nonlinear nature of the estimation. 249 participants

The results regarding gender differences in investing in mixed- and single-sex groups are intriguing—they not only document that men and women differ in the propensity by which they agree to implicit requests for non-desirable tasks, but also provide suggestive insights on what drives these differences. When examining vertical gender segregation, the results from the mixed-sex sessions may become particularly relevant, as gender-specific beliefs may play a central role in environments where women are under-represented.

4. Faculty Volunteering Study

Our laboratory results suggest than men and women differ in their response to tasks that they prefer someone else undertake. In particular we found in mixed-sex groups that women more than men undertake tasks that benefit the collective. To investigate whether this differential extends beyond the laboratory, we obtained data from a naturally occurring volunteer situation.

[38] Interestingly the results are consistent with the predictions of evolutionary game theory, as it predicts in a two player hawk-dove game that a mixed strategy NE will be selected within a single population, whereas a pure strategy equilibrium is predicted when the two players are drawn from different populations (Maynard Smith (1982). Oprea et al (2011) confirm this prediction. Examining investments in a two person Hawk-Dove game they find that play converged to the symmetric mixed Nash equilibrium under a one-population matching protocol, while it moves toward an asymmetric and inequitable pure Nash equilibrium when the participants are assigned either to be row or column players and thus interacted in a two-population environment.

[39] There is no significant gender difference in mean decision time. Clustering the standard errors on the individual a regression of decision time on female and round reveals a coefficient on female of 2.10 (p=0.735) in the mixed-sex sessions and of -6.67 (p=0.272) in the single-sex sessions. Looking at the differential in decision times for women between mixed and single-sex sessions we find an insignificant decrease in decision time, with the coefficient on a single-sex dummy being -3.16 (p=0.595). For men we see instead an insignificant increase in decision time, with the coefficient on a single-sex dummy being 5.591 (p=0.383).

Specifically we analyzed data on volunteering for committees at a large public university. The University in question sends an email each year to all faculty members asking them to serve on a University committee. These committee assignments can largely be viewed as "non-promotable" tasks. While the university benefits from faculty serving on committees, individual benefits are more limited. The assignment takes time away from research and teaching, and plays a negligible role at time of promotion.

Manipulating these email requests Craig Fox and David Tannenbaum conducted an experiment to determine how the language in the email affects the probability that a faculty member agrees to serve on a committee. Fox and Tannenbaum kindly gave us access to their data for the 2012-2013 academic year. As these data contain both the faculty member's response to the email and their demographic characteristics, we can determine whether, when presented with the same task, men and women differ in their willingness to perform the task.

4.1. Participants and Procedure

All faculty members at the university in question were sent an email requesting that they volunteer to join one of several university-wide committees of the Faculty Senate. The email was sent from the Chair of the Faculty Senate to a total of 3,271 faculty members. Our data consists of the faculty member's response and their demographic information (faculty rank, sex, and department affiliation). We coded whether the faculty member was in the medical school (38% were from the medical school) and the faculty member's rank (assistant, associate, full, emeritus, or other).[40] 11.7% were assistant professors, 12.4% were associate professors, 53% were full professors, 21.9% were emeritus professors, and the remaining 1% were categorized as "other" (these included lecturers and instructors). 24.7% of the faculty members were female. Faculty responded to the email in one of three ways: did not respond; declined the request; or accepted the request to join a committee.[41]

4.2. Results

Looking across all faculty 3.7% volunteered to participate on a faculty senate committee, 4.3% responded to the email but indicated that they did not wish to volunteer, and 92% ignored the email. Medical school faculty were more likely to volunteer than non-medical school faculty (5.0% versus 2.6%, $p < .001$). The likelihood of volunteering did not differ significantly between assistant and associate professors, however their rates of volunteering were significantly higher than those of full ($p<.01$) or emeritus professors ($p<.001$).

Turning now to the gender differences in volunteering rates, we see in Table 9 that female faculty were 2.7 times as likely as male faculty to volunteer to be on a committee (7.0% versus 2.6%, $p < .001$).[42]

[40] The small number of volunteers, along with the randomized assignment to experimental treatment, limits the categories that we can examine.

[41] Failure to respond will be treated as declining the request.

[42] Interestingly, one can also examine gender differences in whether or not the faculty member responded to the email if they were not going to volunteer. We think of this as being "polite" and not just ignoring the email: 6.1%

Looking at the results of the probit model for the probability of volunteering, we see in Table 10 that this gender difference is robust to controlling for faculty rank and medical school affiliation.

Table 9: Percent Volunteering to Join a Committee

	All	Male	Female
All (n=3271)	3.7	2.6	7.0***
Medical (n=1233)	5.0	4.1	8.8**
Non-Medical (n=1992)	2.6	1.6	5.4***
Assistant (n=384)	6.5	5.2	8.6
Associate (n=404)	5.7	3.6	9.2*
Full (n=1734)	3.6	2.6	7.0***
Emeritus (n=715)	1.3	1.3	1
Other (n=34)	2.9	0.0	25.0**

*, **, *** gender difference significant at the 5%, 1%, .1% level respectively

Furthermore these gender differences are largely independent of subcategories. The difference remains statistically significant for both medical school and non-medical school faculty. Among the tenure track faculty the gender difference is large and statistically significant for associate professors and full professors. Although the volunteer rate of female assistant professors is larger than the rate for male assistant professors, the difference is not statistically significant (p=.18).

Table 10: Probability of Volunteering to join a Committee (probit)

	Coefficient (p-value)
Female	0.033
	(0.000)
Associate Professor	-0.002
	(0.832)
Full Professor	-0.012
	(0.119)
Emeritus Professor	-0.029
	(0.001)
Other Rank	-0.010
	(0.696)
Medical School	0.023
	(0.000)

Dependent variable: Individual decision to volunteer (1-volunteer, 0-don't volunteer).
The table presents marginal effects. Assistant professor is the excluded category.
n=3,271

of female faculty who did not want to volunteer explicitly declined the request while only 4% of male faculty did so, p<.05.

The volunteer data are consistent with our laboratory findings. In mixed-sex environments, women are more likely than men to undertake tasks that they likely prefer someone do in their place. While previous studies document greater female representation on University committees, the data presented here are particularly informative as they shed light on what gives rise to differential committee participation rates. In documenting differential task assignments it is not possible to determine whether differences result from differences in demand and/or supply. That is assignments may differ because of differences in the frequency by which one is asked to serve on committees, or because of differences in the frequency by which one accepts requests to serve on committees. In looking only at the decision to volunteer our findings demonstrate that differential task assignments can be partially attributed to women being more likely than men to accept requests to perform such non-promotable tasks.

5. Conclusion

A great deal of research has been conducted to help understand gender differences in the labor market. We add to this literature an understanding of the role played by gender differences in the propensity to agree to perform tasks that while helping the group, places the individual performing the task at a relative disadvantage.

When asked to recall requests for undesirable tasks we find that women more than men report finding such requests difficult to decline. In a laboratory study we document that the differential is not limited to different perceptions of such requests. Capturing the incentives members of a group face when asked to find a volunteer for a task that they prefer another member of the group undertakes, our experimental results indicate that women are almost twice as likely to volunteer for such tasks.

To explain the differential we conducted additional sessions using a single-sex rather than a mixed-sex subject pool. Interestingly the gender gap is eliminated when participants know that they are paired only with members of their own sex. Thus group composition plays a central role in explaining why women more than men are willing to undertake the task. In addition to shedding light on why men and women differ in their response to the mixed-sex sessions, the results from the two experiments point to the environment in which we may expect differences in willingness to accept undesirable task assignments to be particularly severe. In particular women may be more likely than men to agree to requests for non-promotable tasks in male-dominated environments.

We interpret the differential response to the single-sex environment as evidence that the initial gender gap results from men and women holding different beliefs about the likelihood that others will invest. An increase in the share of women in the group decreases the likelihood that both men and women invest. This response to gender composition is consistent with the commonly held belief that women more than men will undertake the investment. An alternative explanation for the results is that both men and women are more altruistic toward men. That is both men and women are more generous when the members of their group are more likely to be men. While such an explanation is consistent with women giving less in all-female groups, and consistent with men giving more in all-male groups,

there is no evidence in the literature that suggests that men and women are more altruistically inclined toward men than women. Perhaps the cleanest demonstration of the absence of such an effect is that of Boschini, Muren and Persson (2012). Conducting dictator games in both mixed- and single-sex groups they find, in an environment similar to ours, that the transfers by both men and women are insensitive to the sex of the recipient.

If men and women differ in their propensity to decline requests for non-promotable tasks then they are likely to hold different portfolios of tasks at work. Consistent with our experimental result is the finding female faculty are more likely than male faculty to agree to serve on university committees. In looking at data from a large public university we find that, when asked to volunteer for such an assignment, women faculty are more than twice as likely as men to do so. This suggest that the observed differences in task assignments found in the previous research arise in part from women being more likely than men to accept such assignments. Interestingly the gender difference in accepting 'undesirable' tasks is comparable to what we find in the mixed-sex sessions of laboratory study.

Differences in task allocations can create barriers to the advancement of women in organizations and in society as a whole. By documenting such differences we hope to bring awareness to the effects task allocation may have on the career paths of men and women. While the gender difference in willingness to undertake such tasks is disturbing, is it promising that the differential likely is influenced by beliefs. To the extent that beliefs can be perturbed by awareness, it may be possible to improve the equity of task allocation. Furthermore the finding that volunteering rates differ by gender suggests that institutions may want to reconsider how they allocate undesirable tasks. In acknowledging the differential it may be possible to alter beliefs and the mechanisms we use to assign non-promotable tasks, and this in turn may increase the representation of women in positions of power.

References

Altonji, J., & Blank, R. (1999). Chapter 48 race and gender in the labor market. In Handbook of Labor Economics, Vol 3, Part C, 3143-3259.

Ames, D., Flynn, F.J., & Weber, E. (2004.) It's the thought that counts: On perceiving how favor-gives decide to help, Personality and Social Psychology Bulletin, 30(4), 461-474.

Andreoni, J., & Vesterlund, L. (2001). Which is the fair gender: Gender differences in altruism, Quarterly Journal of Economics, 116(1), 293-312.

Babcock, L., & Laschever, S. (2003). Women don't ask: Negotiation and the gender divide. Princeton, NJ: Princeton University Press.

Baccara, M., Collard-Wexler, A., Felli, L., & Yariv, L. (2012). Child-Adoption Matching: Preferences for Gender and Race, working paper.

Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike? Gender pairing, personality, and cognitive ability as determinants of giving, Journal of Economic Psychology, 25(5), 581-589.

Bertrand, M., Goldin, C., and Katz, L. (2010) Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors, American Economic Journal: Applied Economics, Vol. 2, No. 3, pp. 228-255

Bertrand, M., & Hallock, K. (2001.) The gender gap in top corporate jobs, Industrial and Labor Relations Review, 55(1), 3-21.

Black, S.E., & Strahan, P.E. (2001). The division of spoils: Rent-sharing and discrimination in a regulated industry, American Economic Review, 91(4), 814-831.

Bliss, C., & Nalebuff, B. (1984). Dragon-slaying and ballroom dancing: The private supply of a public good, Journal of Public Economics, 25(1-2), 1-12.

Braiker, H. B. (2001). The Disease to Please: Curing the People-Pleasing Syndrome. New York: McGraw-Hill.

Cadsby, B., Servátka, M., & Song, F. (2009). Gender and generosity: Does degree of anonymity or group gender composition matter?, Experimental Economics, 13(3), 299-308.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences, Journal of Economic Literature, 47(2), 448-474.

De Pater, I., Van Vianen, A., & Bechtoldt, M. (2010). Gender differences in job challenge: A matter of task allocation, Gender, Work, and Organization, 17(4), 433-453.

Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2011). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences, *Journal of the European Economic Association*, 9(3): 522-550.

Duffy, Sean, and John Smith, "Cognitive load in the multi-player prisoner's dilemma game," MPRA Paper No. 30856, May 2011

Eagly, A.H., Wood, W., & Fishbaugh, L. (1981). Sex Differences in Conformity: Surveillance by the Group as a Determinant of Male Nonconformity. Journal of Personality and Social Psychology, 40(2), 384-394.

Eagly, A.H. & Mladinic, A. (1989.) Gender stereotypes and attitudes toward women and men. Personality and Social Psychology Bulletin, 15, 543-558.

Eagly, A.H., & Steffen, V.J. (1984). Gender stereotypes stem from the distribution of women and men into social roles, Journal of Personality and Social Psychology, 46, 735-754.

Eckel, C., & Grossman, P. (1998). Are men less selfish than women: Evidence from dictator experiments, The Economic Journal, 108(448), 726-735.

Eckel, C., & Grossman, P. (2008). Differences in the economic decisions of men and women: Experimental Evidence, Handbook of Experimental Economics Results, 1, 509-519.

Engel, C. (2011). Dictator Games. A Meta-Study, Experimental Economics, 14, 583-610.

Gneezy U, Niederle M, Rustichini A. 2003. Performance in competitive environments: gender differences. Quarterly Journal of Economics. 118:1049–74

Goldin, C., Katz, L., & Kuziemko, I. (2006) The Homecoming of Amercian College Women: The Reversal of the College Gender Gap, Journal of Economic Perspectives, vol. 20, No 4, 133-156.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of 'blind' auditions on female musicians, American Economic Review, 90(4), 715–41.

Heilman, M. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder, Journal of Social Issues, 57, 657-674.

Heilman, M., & Chen, J. (2005). Same behavior, difference consequences: Reactions to men's and women's altruistic citizenship behavior, Journal of Applied Psychology, 90(3), 431-441.

John, O., & Srivastava, S. (1995). The big five trait taxonomy: History, measure, and theoretical perspectives, in Handbook of personality: Theory and research (2nd ed.) (L. Pervin and O. John Eds). New York: Guilford.

Lyness, K.S., & Thompson, D.E. (1997). Above the glass ceiling? A comparison of matched samples of female and male executives, Journal of Applied Psychology, 82(3), 359-375

Major, B., McFarland, D., & Gagnon, D. (1984). Overworked and underpaid: On the nature of gender differences in personal entitlement, Journal of Personality and Social Psychology, 47(6), 1399-1412.

Maynard Smith, John. 1982. Evolution and the Theory of Games. Cambridge University Press.

Mesch, D. J., M.S. Brown, Z.I. Moore, & Hayat, A.D. (2011). Gender differences in charitable giving. *International Journal of Nonprofit and Voluntary Sector Marketing, 16*, 342-355.

Niederle, M., & Vesterlund, L. 2007"Do women shy away from competition? Do men compete too much?" Quarterly Journal of Economics, August 2007, Vol. 122, No. 3: 1067-1101.

Oprea, R. , Henwood, K. and Friedman, D. (2011). "Separating the Hawks from the Doves: Evidence from Continuous Time Laboratory Games." Journal of Economic Theory, 146:6, 2206-2225.

Organ, D.W. (1988). Organizational citizenship behavior: The good soldier syndrome. Lexington, MA: Lexington Books.

Podsakoff, P.M, Mackenzie, S.B., Paine, J.B., & Bachrach, D.G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research, Journal of Management, 26, 513-563.

Polachek, S.W. (1981). Occupational self-selection: A human capital approach to gender differences in occupational structure, The Review of Economics and Statistics, 63(1), 60-69.

Santee, R.T., & Jackson, S.E. (1882). Identity Implications of Conformity: Sex Differences in Normative and Attributional Judgements. Social Psychology Quarterly, 45(2), 121-125.

Schmitt, D. P.; Realo, A.; Voracek, M.; Allik, J. (2008). "Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures". *Journal of Personality and Social Psychology* **94** (1): 168–182.

Weber, E. U., Blais, A., & Betz, E. N. (2002). A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15*, 263-290.

Appendix A: Experimental Instructors and Screen Shot of the Decision Phase

Introduction

Thank you for participating in our study. This is an experiment about decision making. The other people in this room are also participating in the experiment. You must not talk to them or communicate with them in any way. If you have a question please raise your hand and one of us will come to where you are sitting to answer it.

The experiment consists of ten rounds. In each round you are randomly paired with two other participants to form a group. You will never be paired with the same participant twice in a row. Your decisions are anonymous; no one will be able to determine which decisions were made by you. Your round earnings depend on the decisions made by you and by your group members. At the end of the experiment you will be paid, in private and in cash. Your total earnings will equal the sum of your earnings from the ten rounds plus $6 for showing up to the experiment.

Decisions

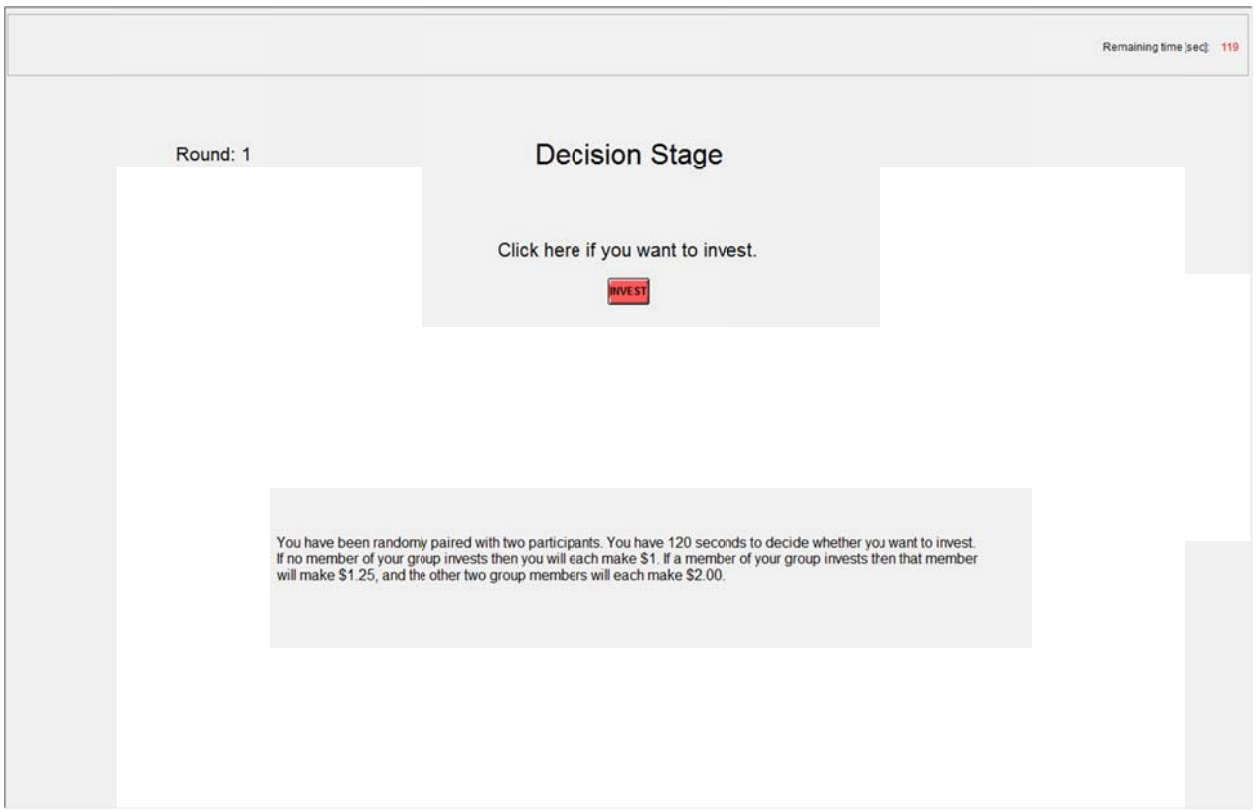
In each round you and the two other group members will have an opportunity to invest in a group account. You and your group members will have 120 seconds to individually decide whether you want to invest in the group account. The round ends when the 120 seconds are up or when the first group member invests in the group account.

Earnings

If no investment is made in the group account, all members of the group will earn $1 for the round. If one group member invests in the group account before the 120 seconds are up, then that individual earns $1.25 for the round and the other two group members each earn $2 for the round. If two group members simultaneously decide to invest, then it is randomly determined which of the two earns $1.25 versus $2 for the round.

Decision Screen

Below you can see a screen shot of the decision screen you will be given to make your investment decision. Listed in the upper right corner is the number of seconds that remain of a round. To the left you can see the round number. The red button in the center of the screen is used to make your investment decision. Please click this button if you wish to invest. The round ends and the decision screen disappears as soon as you or a member of your group invests in the group account.



Summary

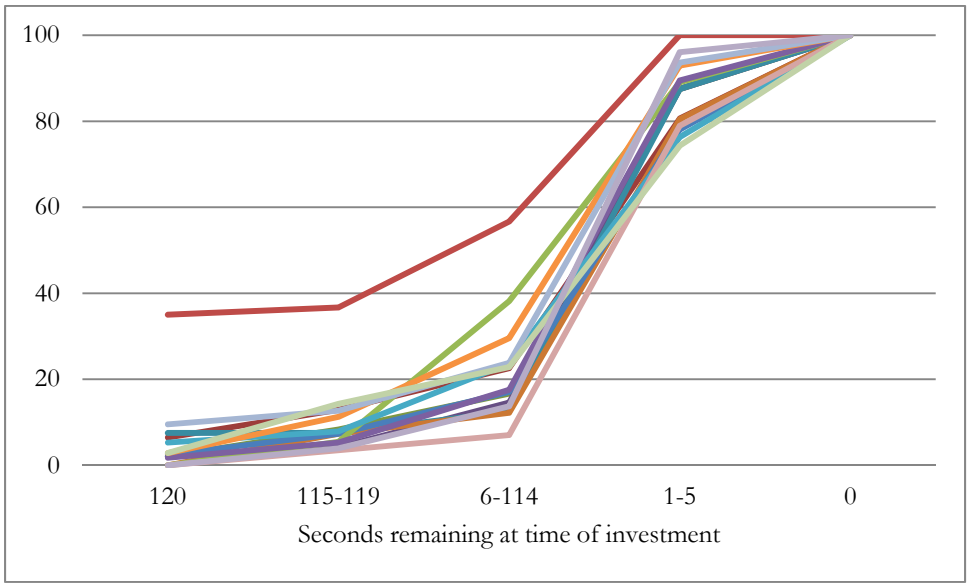
1. In each round you are randomly paired with two other people in this room. You are never paired with the same person twice in a row.
2. A round lasts 120 seconds.
3. During each round you and your group members may invest in the group account. If no one invests you and the two other group members each earn $1 for the round. If one person invests then that person earns $1.25 and the other two group members each earn $2.
4. The round ends once someone invests or when the 120 seconds are up

Please raise your hand if you have any questions before we begin.

Appendix B: Session 7

Examining decision time across the 15 sessions it is transparent that session 7 is an outlier. Participants in this session were more likely to invest in the first second (i.e., when 120 seconds remained of a session) and less likely to invest in the last second of a round. Figure B1 demonstrates for each session the share of investments that were made by the number of second remaining in a round. While the session 7 distribution of decision times differs significantly from the other 14 distributions, none of the other sessions are found to differ from the remaining 14. A two-sample Kolmogorov-Smirnov test for equality of distribution functions yields a p-value of 0.003 when looking at the session 7 distribution, whereas the tests of the remaining distributions differing yields p-values in excess of 0.422 (with the remaining 13 sessions securing p-value in excess of 0.744).

Figure B1 Seconds Remaining at Time of Investment (15 sessions)



The stark differences in decision times suggest that participants in session 7 were playing a different game than that observed in the other sessions, we therefore opted to exclude session 7 from our central analysis. As demonstrated in the following tables the central results of the table are however robust to the inclusion of the data from Session 7.

Table B2: Probability of Investing, n=150, Probit

	All rounds	Round 1-5	6-10
Female	0.111	0.107	0.115
	(0.003)	(0.005)	(0.013)
Round	-0.006	-0.009	-0.009
	(0.056)	(0.409)	(0.349)

Dependent variable: Individual investment decision (1-invest, 0-don't invest).
The table presents marginal effects. Standard errors are clustered on the individual.
p-value is reported in parenthesis. 150 participants

Table B3: Probability of Investing, 150 participants

	(1)	(2)
Female	0.111	0.087
	(0.003)	(0.042)
Round	-0.006	-0.006
	(0.056)	(0.054)
Non-conform		-0.019
		(0.446)
Risk-seeking		-0.024
		(0.236)
Altruism		0.022
		(0.397)
Agreeable		-0.014
		(0.653)

Dependent variable: Individual investment decision (1-invest, 0-don't invest).
The table presents marginal effects. Standard errors are clustered on the individual.
p-value is reported in parenthesis. 150 participants

Table B8: Probability of Investing (probit)

	All rounds	Round 1-5	Round 6-10
Female	0.109	0.106	0.111
	(0.003)	(0.005)	(0.013)
Round	-0.009	-0.011	-0.009
	(0.000)	(0.164)	(0.204)
Single-Sex	0.038	0.054	0.022
	(0.373)	(0.237)	(0.646)
FemalexSingle-Sex	-0.100	-0.097	-0.101
	(0.069)	(0.091)	(0.109)
p-value of test of Female+FemalexSingle=0	0.919	0.916	0.935

Dependent variable: Individual investment decision (1-invest, 0-don't invest). The table presents marginal effects. Standard errors are clustered on the individual. p-value is reported in parenthesis. 267 participants