# Wallflowers Doing Good:
# Field and Lab Evidence of Heterogeneity in Reputation Concerns

**Daniel Jones**[a]

**Sera Linardi**[b]

## Abstract

An extensive literature on reputation signaling has focused on the desire for positive reputation. In our paper we provide field and lab evidence that some individuals are averse to any form of reputation; this aversion correlates with gender in a prosocial setting. We formalize our hypotheses of these "wallflower" types in a theoretical model. The model predicts that wallflowers will deflect unwanted attention by choosing actions that signal that they are an "average altruism type" relative to their audience. Our laboratory experiment supports these predictions.

Keywords: altruism, reputation, signaling, gender, field experiment, lab experiment

[a]Jones: Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA. Email: dbj3@pitt.edu. [b]Corresponding author. Linardi: Graduate School of Public and International Affairs and Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA. Email: linardi@pitt.edu Phone: (412) 648-7650 Fax: (412) 648-7641

## 1. Introduction

How is prosocial behavior impacted by visibility? Given the frequency with which charitable contributions are rewarded with public recognition – be it through a listing of donors in a monthly newsletter or the naming of a building in a donor's honor – fundraisers seem to believe that visibility has a positive effect. Yet existing research has arrived at ambiguous and sometimes contradictory conclusions. Some researchers have provided evidence that visibility increases voluntary contributions to a public good (e.g, Andreoni and Petrie (2004), Rege and Telle (2004)). On the other hand, Dufwenberg and Muren (2006) find that visibility *decreases* giving in a dictator game. Others have found that the success of visibility depends heavily on the context or nature of the decisions being made (Gachter and Fehr (1999), Soetevent (2005), Alpizar et al. (2008), Shi (2011)).[1] With divergent results across relatively similar environments, it remains unclear how visibility impacts prosocial behavior.

In our paper, we provide evidence that 1) there is heterogeneity in reputation concerns; specifically, some individuals are not comfortable signaling their altruism, and 2) the impact of visibility on prosocial behavior depends on reputation concerns. Prosocial actions are fraught because they are revealing: purely voluntary actions are relatively accurate signals about types. When a person averse to reputation signaling is forced to choose among actions with varying degrees of signaling content, she may respond by choosing the action that has the least signaling content. This strategy often corresponds to conforming to what she perceives as the norm. One example is Linardi and McConnell's (2011) experiment where participants volunteer for a nonprofit. When the first individual quits volunteering and leaves the session, a large portion of

[1] Gachter and Fehr (1999) find that contributions to a public good increase with visibility when contributors meet and interact beforehand, but not otherwise. Soetevent (2005) in a field experiment manipulates the visibility of church donations and finds a positive impact for external causes but no impact of internal fundraising. Alpizar et al. (2008) conduct a field experiment where they approach visitors in a national park and solicit donations; allowing solicitors to directly view visitors' choices causes a small increase in contributions. Shi's (2011) survey results show the opposite: visibility alone (*without* monetary or other material rewards) decreases survey respondents' hypothetical willingness to donate blood.

the remaining participants immediately follow suit. The propensity toward norm conformance was not observed in treatments in which quitting did not signal lack of altruism.[2]

In our field experiment, carnival attendees on a college campus can spend as much time as they like at a fundraising booth completing word search puzzles; each word they find generates a small contribution for a charity. In one treatment participants observe the names and total contributions of others who have participated before them and are aware that future participants will observe their name and contribution. When contributions are publicized, males give more while females give less. This is because females, who tend to complete multiple puzzles in the private treatment, respond to higher visibility by following the norm of completing exactly one word search puzzle.[3]

The field experiment motivates us to propose a new theoretical model, which we then test with a laboratory experiment. Benabou and Tirole's (2006) model of prosocial behavior posits that honor (or signaling a type above the average) provides positive utility while stigma (signaling a type below the average) provides negative utility.[4] We depart from this model by assuming that some individuals act like "wallflowers". *Any* reputation, whether positive or negative, brings them unwanted attention.[5] Therefore, wallflower types choose actions to minimize any inferences that can be drawn from their behavior.

Our field results suggest that with regards to prosocial behavior, gender is an important predictor of reputation concerns. Females are more likely to be "wallflower" types – perhaps because of cultural conditioning or other correlated personality differences – and males are more likely to hold preferences consistent with the standard Benabou and Tirole model. Thus, our model

[2] No cascade was observed when there was a possibility that individuals may have quit due to computer generated time limits on volunteering.
[3] When we refer to "norms" both here and throughout the paper, we are referring to "descriptive norms" or *commonly available information about the typical behavior of others* (following the definitions of Croson et al. (2009) and Cialdini et al. (1990)) as opposed to "injunctive norms" (informal rules about "appropriate" behavior.)
[4] Experimental evidence on Benabou and Tirole (2006) model shows that monetary behavior only has a negative effect on volunteering if it is offered in public setting, where a volunteer's prosocial behavior also serves as a signal of his altruism (e.g Ariely et al. (2009), Lacetera and Macis (2010), Linardi and McConnell (2008)).
[5] Within the context of Benabou and Tirole (2006) the marginal value of publicizing prosocial behavior decreases when it causes observers to suspect contributors to be reputation-seeking. However, this will not result in a decrease in contribution or in higher norm conformance as in our model.

predicts that visibility encourages males to increase their contributions while inducing females to minimize the distance between their contributions and others' expected contribution. Whether this will result in an increase or decrease in average female contribution will depend on the location of the norm.

We test these predictions by designing a laboratory experiment where we can observe whether subjects' contributions respond to a set of possible norms. Upon arrival to the lab, a computer interface assigns subjects to anonymous groups of three. In the first stage, subjects decide how much to contribute to a nonprofit. Participants in the *Baseline* treatment are informed that their donations will be submitted in a sealed envelope at the end of the experiment. Subjects in the *Visibility* treatment are told that they will meet their group members at the end of the experiment to submit their donations in front of one another. In the second stage, subjects are given an opportunity to change their donation by conditioning their contribution on every possible combination of the other two group members' donations. We find support for our predictions: females in the Visibility treatment are more likely to condition their donations to fall between their group members' contribution, thus signaling that they are "average types" in relation to their group members.[6]

The fact that "wallflower" behavior is more common amongst females is consistent with a long line of literature on gender differences. Croson and Gneezy (2009), in a review of research from psychology and economics, suggest, "the social preferences of women are more situationally specific than those of men." A number of recent studies provide further evidence of this claim (Zetland and Della Giusta (2011), Mellstrom and Johannesson (2008), Lacetera and Macis (2010)).[7] It has also been shown that females tend to avoid competitive situations altogether when given the choice (Niederle & Vesterlund, 2007). Additionally, Kanthak and Woon (2012)

[6] Although not focused on gender, Yariv and Goeree (2007) also found conformity in lab subjects' frequent choice of statistically uninformative sequences of play from predecessors instead of informative private draws.

[7] Zetland and Della Giusta (2011) vary the salience of social information in a public goods game and show that only women change their behavior in response to this manipulation. Others have shown that females are significantly less likely to donate blood when a monetary payment is offered, while males are more likely to donate (Mellstrom and Johannesson (2008), Lacetera and Macis (2010)). This may be due to females internalizing social cues that blood donation should be given out of pure altruism and not for money. There are no observable gender differences in crowding out when extrinsic incentives are given in the form of a gift voucher (Lacetera and Macis (2010)) or when participants are allowed to donate the cash payment to charity (Mellstrom and Johannesson (2008)).

show that females are reluctant to reveal their ability to compete in an election environment. If publication of prosocial behavior is perceived as a status competition, it is unsurprising to find that females "opt out" of this competition by choosing actions that are least likely to draw attention.

More generally, the heterogeneity in reputation concerns uncovered here may explain the divergent results in the literature. Consider the contrasting results of Rege and Telle (2004) – where visibility increases giving – and Dufwenberg and Muren (2005) – where visibility decreases giving. In Rege and Telle's public goods game with visibility, a majority of participants choose the socially efficient action, contributing their entire endowment. Given this norm of high giving we would predict the same high contribution from both the reputation-seeking and wallflower types. However, while the former are motivated by altruism signaling, the latter do so to conform. On the other hand, Dufwenberg and Muren's dictator game typically has a norm of equal division (Andreoni and Bernheim, 2009). In this setting, visibility will induce higher contributions from the reputation-seeking types and more equal divisions from the wallflowers types. Indeed, the decrease in average contributions in their visibility treatment seems to be partially driven by a decrease in the frequency of choosing a contribution higher than an equal split, and this is particularly true amongst females. Our results suggest that it may not be appropriate to ask whether visibility increases or decreases prosocial behavior, as the impact could depend heavily on whether the relevant norm dictates high giving, low giving, or equal division. The success of public recognition should ultimately be expected to depend on the location of the norm and the heterogeneity of reputation concerns in the population of interest.

Our evidence on the gender differences in signaling prosociality has important practical implications. The fundraising community has long debated whether males and females differ in their response to public recognition. Some argue that "women do not like solicitations based on peer pressure, competition or public recognition" (Taylor & Kaminski, 1997), while others claim that there is little evidence for this long-held belief (Hall, 2004). Our results lend credence to the former argument, suggesting that nonprofits that rely heavily on contributions from women should be cautious in implementing donor recognition programs.

The rest of the paper proceeds as follows. In Section 2, we discuss the design and results of our field experiment. In Section 3, we present a simple model to capture our observation from the field. In Section 4, we describe the design of our laboratory experiment. An analysis of the laboratory results is presented in Section 5. Section 6 concludes.

## 2. Field Experiment

### 2.1 Experimental Design

Increasing the visibility of prosocial actions can affect behavior through multiple channels. An individual may be more concerned about her reputation since her actions now carry a signal about her altruism. That same individual may also learn more about the benefits, costs, and expectations regarding the prosocial act through greater publicity of other's contribution. Our field experiment was initially designed to disentangle the first mechanism (reputation concerns) from the second (social information).

The setting of our field experiment is the California Institute of Technology annual campus carnival, which is regularly attended by about 1500 people. During carnival, approximately 40 student clubs are each assigned a 10x10 foot booth to advertise their group's activities. We partnered with the Caltech chapter of Engineers for a Sustainable World (ESW-CIT) to raise money for a water project in Honduras during the event.[8] The festive and social atmosphere of the carnival provides us with a natural opportunity to simulate the environment in which fundraising often takes place in.

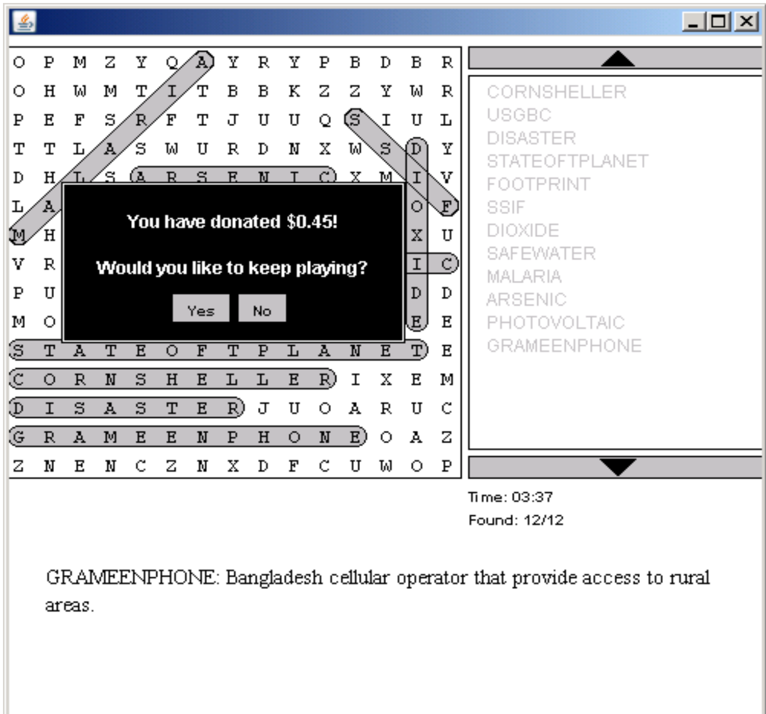
The fundraising task utilized a computer word search game titled *Find Words, Give Water*.[9] In this game, a 15x15 grid of jumbled letters hides twelve (12) words related to environmental sustainability. The computer screen displays an alphabet grid and a list of sustainability terms hidden within that grid (Figure 1). Words are found by highlighting the letters that make up the word. For every word found, 15 cents are donated to a water project in Honduras. Players can play for as long as they like and they can quit at any point by clicking on a button that says *End*

[8] ESW is a national organization of students engaged in social activism through collaborative technical projects.
[9] This game is written in the JAVA programming language. Code is available upon request.

*Game*. When this button is clicked or upon participants' completion of a puzzle, a pop-up box displays participants' current donation and asks the participants if they wish to continue. Clicking on *Yes* returns them to the puzzle. Clicking on *No* brings them to a survey.

**FIGURE 1**

**A word search puzzle**




The experiment is conducted on five computers that were set up in the ESW booth. Participants playing the game are first greeted by a screen (Figure A (Appendix) 1) that explains the game and the project. When participants click a button that says *Proceed,* the software randomly assigns the participants into one of three treatments: Baseline, Scores, or Names.

In the Baseline treatment, participants immediately proceed to the word search game, as described above. Participants in the Scores treatment are first shown a screen that lists the last 10 donations on the computer terminal where they are seated (Figure A2). These donations include all participants who were not asked to identify themselves (that is, participants in Baseline and Scores treatments). After viewing previous donations, participants click *Start* and proceed with the game.

Participants in the Names treatment are asked for their full names after advancing from the first page (Figure A3). They are then shown a screen that lists the full names and donations of the last 10 participants on the computer terminal (Figure A4). These donations include only participants in the Names treatment, since participants from the other two treatments did not enter their names. After clicking *Start*, participants are directed to the word search game and proceed just as participants in the Baseline and Scores treatments.

Our hypotheses are the following: The Scores treatment will increase contributions relative to the Baseline through social information about what others are giving. The Names treatment will increase contribution relative to the Scores treatment through the additional impact of reputation signaling.

After participants decide to stop working, they are directed to a brief survey (Figure A5) in which they are asked to indicate their gender, academic standing, major, and whether they are interested in being contacted about volunteer opportunities in the future. These survey responses are summarized in Appendix Table 1.

## 2.2 Results

We begin by documenting some basic patterns observed in the data. On average, participants spend 5.5 minutes working and contribute $2.25, which corresponds to finding roughly 15 words. Perhaps not surprisingly, participants tend to stop at the end of a word search puzzle – this is the case for 70 of the 104 participants. Roughly half of all participants stop at the end of the *first* puzzle, finding exactly 12 words and contributing $1.80. This distribution of donations is illustrated in Figure A6.

Table 1 provides an initial comparison of the effect of each treatment on donations. Overall, 104 individuals participated; this includes 29 women, 46 males, and 29 people who failed to indicate their gender on the survey. Splitting treatments by gender results in very small sample sizes;

therefore we take the field results only as suggestive evidence, which we test more systematically in our laboratory experiment.

**TABLE 1**

**Mean donation by treatment**

| | All | Male | Female | No gender indicated |
|---|---|---|---|---|
| Baseline | 1.95 | 2.03 | 2.43 | 1.56 |
| | (0.191) | (0.237) | (0.451) | (0.408) |
| | n=34 | n=19 | n=5 | n=10 |
| Scores | 2.57 | 2.79 | 3.15 | 1.58 |
| | (0.351) | (0.607) | (0.741) | (0.189) |
| | n=40 | n=17 | n=12 | n=12 |
| Names | 2.18 | 2.94 | 2.08 | 1.37 |
| | (0.332) | (0.900) | (0.248) | (0.246) |
| | n=30 | n=10 | n=12 | n=8 |
| Scores - Baseline | 0.62 | 0.76* | 0.72 | 0.02 |
| *p-value* | *0.19* | *0.07* | *0.96* | *0.81* |
| Names - Scores | -0.39 | 0.15 | -1.07 | -0.21 |
| *p-value* | *0.57* | *0.84* | *0.64* | *0.71* |

*Reported p-values are from Wilcoxon tests*

Table 1 provides some indication that social information (Scores) increases donations relative to the baseline. The marginal impact of associating names to donations (Names) is not significant, but the sign is surprisingly negative for females. Although only there is just one significant treatment effect (Scores treatment for males), there is reason to believe that these simple means comparisons mask the full impact of the treatments. Given that the computers are not networked, Names and Scores participants observe only the previous scores from the computer at which they are seated; therefore, treatment impact may vary across computer terminals. Additionally, the number of people in the booth fluctuates throughout the day; this is relevant given that Linardi & McConnell (2011) find that participants work longer on a volunteering task as the size of the peer audience increases. In Table 2, we present the results of several models in which we regress total donation on treatment indicators in an attempt to address these issues. Columns 2 through 4 include computer fixed effects to account for computer-specific display of previous scores. Standard errors are clustered for each computer-treatment group to account for the correlation of errors across observations within the same treatment and computer. Columns 3 and 4 include "booth count" fixed effects to account for the number of people present in the booth (a number between 0 and 4) at the time that the subject leaves. The dummy variable *Finished* accounts for the fact that ¾ of the participants stop at the end of a word search puzzle (thus contributing in multiples of $1.80).

**TABLE 2**

**Regression estimation results -- dep. var.: donation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Scores | 0.61 | 0.79** | 1.28** | 1.32** |
| | (0.45) | (0.36) | (0.51) | (0.52) |
| Names | 0.22 | 0.40 | 0.63 | 1.58 |
| | (0.39) | (0.38) | (0.57) | (0.99) |
| Finished | | 0.82** | 0.59 | 0.66 |
| | | (0.38) | (0.46) | (0.49) |
| Female | | | -0.50 | 0.53* |
| | | | (0.55) | (0.30) |
| Scores*Female | | | | -0.63 |
| | | | | (0.89) |
| Names*Female | | | | -2.40** |
| | | | | (1.06) |
| Constant | 1.95*** | 1.30*** | 1.98*** | 1.77*** |
| | (0.20) | (0.42) | (0.47) | (0.53) |
| Computer Fes | | X | X | X |
| Booth count Fes | | | X | X |
| R-squared | 0.02 | 0.14 | 0.20 | 0.24 |
| N | 104 | 104 | 75 | 75 |

Standard errors (clustered for each computer-treatment group) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Column 1 presents a simple baseline model with no controls. Consistent with the simple means comparisons from above, both the Scores and Names treatments appear to result in higher contributions than the Baseline, but these differences are not significant. Upon controlling for computer fixed effects in model 2 and gender and booth count in model 3, the positive impact of score revelation becomes clear. To explore gender-specific treatment effects, model 4 includes an interaction of treatment and gender.[10] *Names * Female* is significant and negative thus confirming that revelation of names decreases giving among females. Moreover, the sum of the

[10] Note that the number of observations drops to 75 in columns 3 and 4, as we omit the 29 participants who did not identify their gender.

coefficients *Names* and *Names * Female* – the treatment effect conditional on being female – is significantly different than zero. Thus, although it appears that in the aggregate the Names treatment has a weakly positive effect on giving, it in fact has a strong impact that varies significantly across genders.

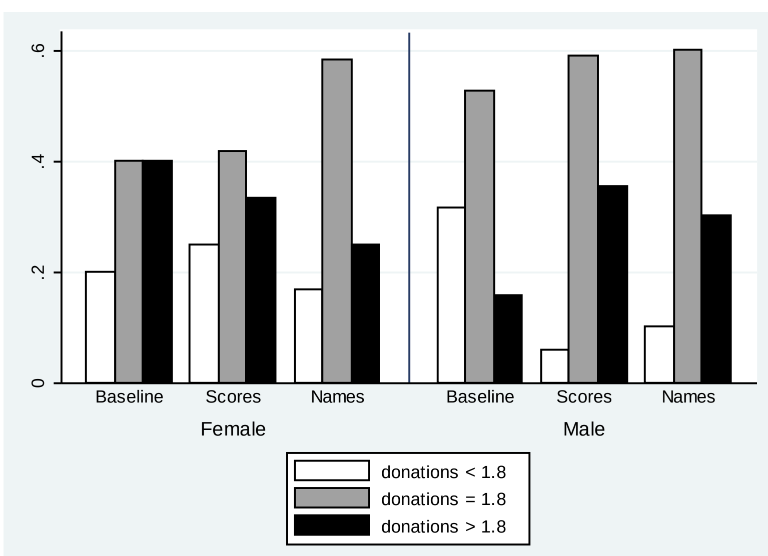
**Field result 1:** Information about previous contributions increases contribution regardless of gender.

**Field result 2:** Conditional on receiving information about previous contributions, revelation of identity has a negative impact on female contributions.

Why does revelation of names decrease female giving? Recall that half of the participants stop after completing one word search puzzle, making $1.80 the modal and median donation amount. Figure 2 below displays the frequency with which participants choose contributions below, equal to, or greater than $1.80. It suggests that the decrease in female giving is not driven by a downward shift in the distribution of their contributions; instead, it is driven by the higher frequency with which females follow the $1.80 contribution norm.

**FIGURE 2**

**Freq. of donations below, equal to, and above $1.80 by treatment and gender**



The propensity to give $1.80 is evident in a linear probability model (Table 3) in which we estimate the likelihood that an individual chooses a contribution of $1.80 conditional on their gender and treatment.[11] The positive and significant coefficient on *Names * Female* in Column 2 confirms the message of Figure 2: females are much more likely to conform to $1.80 norm in the Names treatment.

**TABLE 3**

**Likelihood of contributing $1.80**

| | (1) | (2) | (3) |
|---|---|---|---|
| Scores | 0.0000 | -0.0110 | 0.481 |
| | (0.0864) | (0.114) | (0.333) |
| Names | 0.0764 | -0.124 | -0.147 |
| | (0.117) | (0.201) | (0.298) |
| Female | -0.0402 | -0.263** | -0.201 |
| | (0.107) | (0.109) | (0.172) |
| Female*Scores | | 0.144 | -0.308 |
| | | (0.149) | (0.420) |
| Female*Names | | 0.508** | 0.269 |
| | | (0.232) | (0.559) |
| Constant | 0.627*** | 0.682*** | 0.799*** |
| | (0.135) | (0.164) | (0.232) |
| Computer Fes | X | X | X |
| Booth count Fes | X | X | X |
| *Median observation not $1.80* | | | X |
| Observations | 75 | 75 | 42 |
| R-squared | 0.144 | 0.174 | 0.168 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We include a robustness check for norm conformance in Column 3 of Table 3, which estimates the same equation as in column 2, but excludes participants whom observed a median contribution of $1.80. Our sample now consists only of participants who observe a set of contributions with some other median, participants who observe nothing because they are the first to sit at their computer, and Baseline treatment participants. If females are simply more likely to choose a contribution of $1.80 in Names for some reason other than observing that

[11] A linear probability model is used for ease of interpretation. A probit estimation provides similar results.

$1.80 has emerged as the norm, then the results from Column 2 should hold with this exclusion. However, *Names * Female* is not significant in Column 3.

What sort of preferences might explain females' increased propensity to follow the $1.80 norm in the Names treatment? One possibility is that females may have perceived the revelation of identity as opening up their actions to unwanted scrutiny. In response, they attempt to "hide" by choosing an action that allows them to be unnoticed; following the $1.80 norm is optimal in this environment since it signals that an individual neither too selfish nor too altruistic relative to everyone else.[12] In the next section, we present a simple model to capture this argument with greater precision.

## 3. Model

Our model builds on the honor/stigma version of Benabou & Tirole's (2006) model of prosocial behavior (hereafter BT). This version of the model is particularly well suited to our argument as it highlights how agents might respond to norms while allowing these norms to be endogenous. Agents are (to varying degrees) motivated to engage in prosocial behavior, but are also concerned about how their actions will be perceived by others. Unlike BT, whose model is driven by honor-seeking and stigma-avoidance, we assume that some individuals – who we refer to as "wallflower" types – avoid both honor *and* stigma, and therefore prefer to choose the action least likely to draw attention.

An individual's type is described by $(v,g)$ where $v \sim u[0,A]$ denotes intrinsic altruism and $g \in$ {*Reputation-seeking*, *Wallflower*} indicates reputation concerns. We denote average altruism as $\bar{v} = A/2$ – note that the average here coincides with the median due to the symmetric distribution. Individuals choose between three contributions levels: low, medium, or high, where each

[12] Alternatively, the list of scores shown before play contains additional information in the form of names of previous players. Females might be more persuaded by social information when identities are attached to choices, and thus chose $1.80 more often. Column 3 provides some suggestive evidence in support of the first explanation, but does not completely eliminate the confounding factor of stronger social information. This concern is eliminated in our laboratory experiment.

contribution level has increasing marginal costs.[13] For simplicity, we will represent the contribution levels as $a \in \{0,1,2\}$ where $c(a)=ka^2$. We will write $c_1$ for $c(1)$ and $c_2$ for $c(2)$ as a shorthand.

An individual's utility function is:

$$u(a|g,x,v) = va - c(a) + xR(a|g) \quad (1)$$

where x=1 if the contribution is visible and 0 otherwise, and $R(a|g)$ represents the reputational benefit when contributions are visible.

An individual experiences honor when her expected altruism, given her contribution level, is higher than the average type, or $E(v|a,g) - \bar{v} > 0$. Conversely, an individual experiences stigma when her expected altruism is lower than average type, or $\bar{v} - E(v|a,g)>0$.

$$R(a|g) = h(g)max[E(v|a,g)- \bar{v},0]\text{-max } [\bar{v} - E(v|a,g),0] \quad (2)$$

For *Reputation-seeking* individuals, honor yields positive utility while stigma yields negative utility. This corresponds directly to BT’s model. *Wallflower* individuals, however, are uncomfortable with any type of reputation; hence their utility decreases with both stigma and honor.

$$h(g)=1 \quad \text{if} \quad g = M \text{ (Reputation-seeking)}, \quad (3)$$

$$h(g)=-1 \quad \text{if} \quad g = F \text{ (Wallflower)}$$

Our field experiment suggests that in a prosocial setting, heterogeneity in reputation concerns is correlated with gender. For the remainder of the paper we make the simplifying assumption that this correlation is perfect; every female is a wallflower type and every male is not. The comparative static predictions that result are the same as if we had made the more realistic assumption that females are simply more likely than males to be wallflowers (but are not guaranteed to be wallflowers). We therefore adopt a short hand of M for Reputation-seeking types and F for Wallflower types.

[13] Because any portion of the endowment that is not contributed is consumed, this assumption can also be viewed as decreasing marginal benefit of consumption.

Define $v_a^g$ and $\tilde{v}_a^g$ as the cutoff altruism type of an individual with reputational concern of type *g* who is indifferent between contributing *a* or *a*-1 when contributions are not visible and visible, respectively. We restrict our attention to the range of costs where, without visibility, the probability that an individual will contribute at the high level is small but positive and where $v_1^g \leq \tilde{v}_2^g$.[14]

When there is no visibility, reputation concerns do not play any role since actions do not generate any signals. Cutoff types are solely defined by the marginal cost of effort $\Delta c(a) \equiv c(a)-c(a-1)$.

$$v_a \equiv v_a^M = v_a^F = \Delta c(a) \quad (4)$$

Reputation concerns matter when contribution levels are linked to an individual's identity and publicized. For *Reputation-seeking* types, reputation benefit increases in the level of altruism signaled. Therefore, marginal reputation benefit at any level of effort is always positive ($r(a|M)>0$). Increasing visibility has the effect of decreasing the level altruism necessary such that a *Reputation-seeking* individual becomes indifferent between contributing a and a-1.

$$\tilde{v}_a^M = \Delta c(a) - r(a|M) < v_a^M \quad (5)$$

For *Wallflower* types, the further the type signaled by an action is from the average type, the less reputation benefit there is in taking that action. Therefore, marginal reputation benefit is positive for actions that bring an individual's expected type toward the average and negative for actions that do otherwise.

$$\tilde{v}_a^F = \Delta c(a) - r(a|F) > v_a^F \text{ when } E[v|a-1; F; x] > \bar{v} \quad (6)$$

and

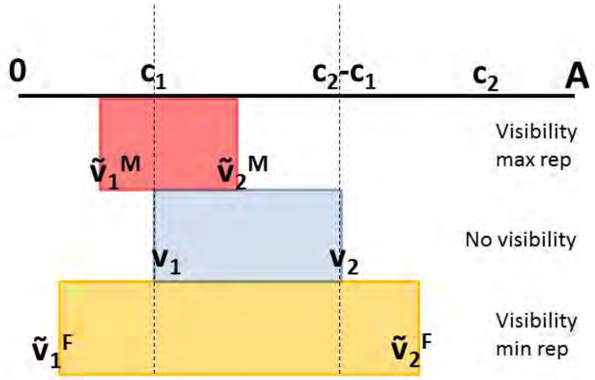
$$\tilde{v}_a^F = \Delta c(a) - r(a|F) < v_a^F \text{ when } E[v|a; F; x] < \bar{v}$$

[14] See Theoretical Appendix (Section 1) for details and proofs.

Figure 3 below illustrates the intuition above, with shaded areas representing the range of values that can be taken by the altruism variable *v* for individuals who choose to contribute at the medium level.

**FIGURE 3**

**Changes in cutoff types due to increased visibility as a function of reputation concerns**




The impact of publicizing previously invisible prosocial behavior is summarized in Theorem 1, below. All derivations are provided in Appendix B.

**Theorem 1:** Cutoffs types satisfy the inequalities below:

$\tilde{v}_1^F < \tilde{v}_1^M < v_1 < \tilde{v}_2^M < \bar{v} < v_2 < \tilde{v}_2^F$

where $v_1 = v_1^M = v_1^F = c_1$

and $v_2 = v_2^M = v_2^F = c_2 - c_1$

**Corollary:**

(i) Visibility increases the proportion of *Wallflower* types that choose the middle action.

(ii) Visibility increases the proportion of *Reputation-seeking* types that choose the high action.

The model does not provide clear predictions on the impact of visibility on average contribution of *Wallflower* types and on the dispersion of contribution among Reputation-seeking types. We can however, state the two propositions below.

**Proposition 1**

Visibility increases average contribution of *Reputation-seeking* types.

**Proposition 2**

Visibility decreases dispersion of contribution among *Wallflower* types.

## 4. Laboratory Experiment

We now turn to the design and results of our laboratory experiment. To test our hypothesis, it would be ideal to manipulate a norm in a charitable giving task and observe whether males and females show different tendency to conform to the norm as visibility is increased. However, “manipulating norms” in the laboratory is not a straightforward task. We therefore assign subjects to groups and use the strategy method to elicit subjects’ giving strategy *conditional on the contributions of their group members* who, in one of the treatments, will observe what they gave.[15] By doing so, we are able to observe what a particular subject would choose to do under a variety of different norms.

### 4.1 Experimental Design

Subjects were recruited through the Pittsburgh Experimental Economics Laboratory (PEEL) database. In each session, fifteen subjects were seated at computer terminals upon arrival and randomly assigned to groups of three and identified only by anonymous subject IDs. We explained that the experiment would consist of a “giving task,” during which they would have the opportunity to donate to a charitable cause, and a “guessing task,” during which they would have the opportunity to earn “up to an additional $7.”

The software then played a slideshow of a water project in Tingo Pucara, Ecuador, organized by Engineers Without Borders (EWB) Pittsburgh, our partner nonprofit for the laboratory

[15] The conditional contribution elicitation is based on the design of Fischbacher et al. (2001), but with some important differences. Fischbacher et al. allow participants to condition on the mean of their group members’ contributions whereas we allow participants to condition on every possible combination of group members’ contributions.

experiment. After the slide show, we endowed participants with 10 one-dollar bills in an envelope, from which they would make their contributions. Any money that participants did not donate was theirs' to keep.

There are two treatments: *Control* and *Visibility*. The only difference between these two treatments is that *Visibility* treatment subjects know that their identity will be revealed to their group members and the experimenter after all decisions have been made.

In the control treatment, the experimenter read from the following script:

> *"At the end of today's session, you will leave your donation in its original envelope on your desk. The software will inform you of your group's total donation to Tingo Pucara."*

In the visibility treatment, the experimenter read from the following script:

> *"At the end of today's session, you and your group will sit down together around a table to submit your contributions. You will go to a different room with the experimenter who will then collect each group member's contributions, announce how much each person gave, and announce the total donation to Tingo Pucara. Your group members are the only participants who will observe how much you chose to give."*

After these preliminaries were completed, the decision-making portion of the experiment consisted of three phases: (1) unconditional contributions, (2) conditional contributions, and (3) belief elicitation. These decision tasks are explained in detail below. Participants did not learn of the details of any of these phases until they occurred nor were participants aware that the "giving task" would ultimately consist of both an unconditional and conditional contribution task.

We conducted five sessions of each of the two treatments with 15 participants per session. There were (roughly) equal proportions of males and females in all ten sessions, with 75 males and 75 females participating overall. All sessions were conducted in the Pittsburgh Experimental

Economics Laboratory (PEEL) using z-Tree software. Subjects received a $5 show-up fee in addition to any money kept or earned during the experiment. Sessions lasted less than one hour.

#### *4.1.1 Unconditional contribution*

In the *unconditional contribution* task, participants were simply asked to indicate how much of their $10 endowment they wanted to donate. We restricted contributions to multiples of 2; that is, participants could choose to give $0, $2, $4, $6, $8, or $10. This restriction and the small group sizes were chosen to limit the number of choices that participants would face in the conditional contribution task.

#### *4.1.2 Conditional contribution*

In the *conditional contribution* task, participants were given the opportunity to change their contribution based on what the other two members of their group chose. Participants choose contributions conditional on *every possible combination* of their group members' contributions from the unconditional phase – a total of 21 decisions. For instance, in the first screen (Figure 4), a subject is asked to assume that one of their group members gave $0 in the unconditional phase. A list of all possible unconditional contributions of the second member of their group ($0, $2, $4, $6, $8 and $10) is displayed, and the subject is asked to indicate her donation for each combination of hypothetical contributions. A similar series of screens then follows. In the second screen, subjects are asked to assume one group member gave $2 and to then indicate how she would contribute if the second group member gave $2, $4, $6, $8, or $10. The following screens present the rest of the scenarios, fixing one group member's contribution at $4, $6, $8 and $10.

As in the conditional contribution design of Fischbacher et al. (2001), one member of each group was randomly selected at the end of the session to have her conditional contribution implemented. Thus, when participants submitted their contributions at the end of the experiment, two members of each group submitted the contribution they chose in the unconditional phase and the remaining member submitted the relevant conditional contribution. Even in the visibility treatment, participants never learned which member of their group was randomly selected to have his or her conditional contribution implemented.

**FIGURE 4**

**Conditional contribution entry**


### *4.1.3 Belief elicitation*

In the *belief elicitation* task, participants were asked to guess the number of people who chose each of the possible unconditional contributions. Participants were informed that they had 14 tokens (one for each of the other participants in the room) to allocate across the possible unconditional contributions ($0, $2, $4, $6, $8, and $10). Each token that was placed correctly earned the participant $0.50. Denoting a participant's *reported guess* of the number of subjects who chose unconditional contribution $k$ as $g_k$ and the actual number of subjects as $n_k$, then the belief elicitation payoff can be expressed as:

$$\sum_k 0.5 * \min\{g_k, n_k\} \;\; for\; k \in \{0, 2, 4, 6, 8, 10\} \qquad (7)$$

We show that this belief elicitation task is incentive compatible in Appendix C.

After completing the belief elicitation task, participants completed a brief survey where they indicated their gender and familiarity with the charitable cause.[16] They were next informed of (1) their earnings from the belief elicitation task, (2) the actual contribution they would provide based on whether or not they were the randomly selected member of their group and, if so, the unconditional contributions of their group members, and (3) their group's total contribution to the cause. Donations were then collected according to the procedures described above.

## 4.2 Results

Our laboratory experiment investigates whether systematic differences exist in the manner in which individuals respond to the revelation of their identity. We focus on how perceived norms, namely individuals' beliefs about others' behavior, influence the choices of female and male subjects when these choices will be revealed to an audience.

The experimental design allows the influence of perceived norms to be investigated in two contexts. First, the relationship between unconditional contributions and elicited beliefs reveals how subjects make decisions when they are uncertainty about their audience. Second, conditional contributions provide subjects' complete giving strategy for all hypothetical audiences. Since the audience for an individual's decision is her two group members (drawn randomly from all session participants), elicited belief can be interpreted as her subjective probability distribution over possible audience types. This belief links a subject's choices across the two contexts.

We will test three hypotheses from the theoretical predictions. In Section 4.2.1, we test Propositions 1 and 2 through simple means comparisons and regression analyses. We find that visibility increases male giving and decreases the variance of female giving. In subsection 4.2.2, we test whether the results in Section 4.2.1 are driven by the mechanism described in Theorem 1.

### *4.2.1 Male Mean (Proposition 1) and Female Variance (Proposition 2)*

[16] There is no evidence of a gender difference in familiarity to the organization (EWB). Nationally, 45% of EWB members are women.

We begin our analysis with simple comparisons of means. Throughout, let $c_i^u$ denote i's unconditional contribution and let $c_i^c$ denote i's conditional contribution. Table 4 reports the impact that visibility has on four outcomes: unconditional contributions ($c_i^u$), expectations about others' unconditional contributions constructed from the belief elicitation task ($E_i[c_{-i}^u]$), "belief-normalized" unconditional contributions ($c_i^u - E_i[c_{-i}^u]$), and conditional contributions ($c_i^c$). In Panel D conditional contributions have been aggregated to the individual-level; we first average each subject's 21 conditional contributions before averaging across all subjects.

**TABLE 4**
**Means -- Initial assessment of treatment effects**

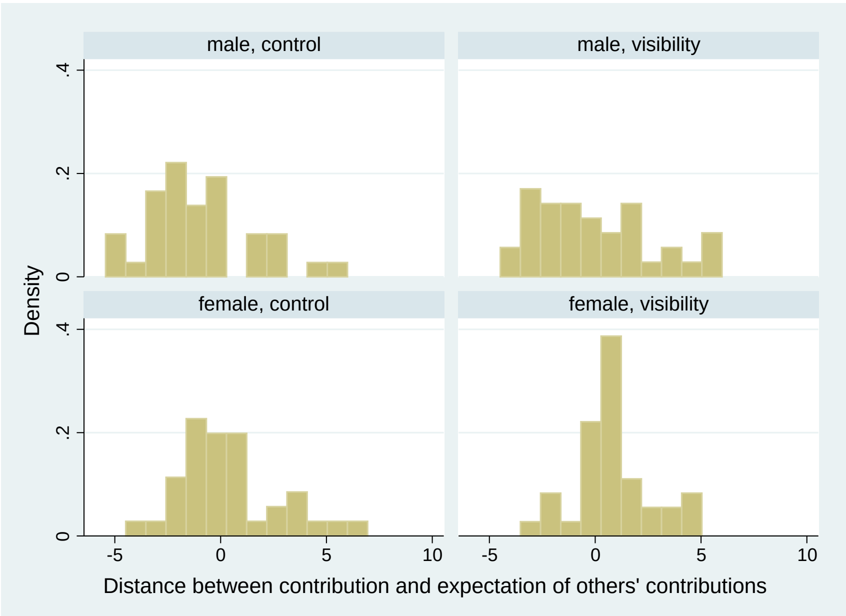
| | | Control | Sig. diff. | Visibility |
|---|---|---|---|---|
| (A) Unconditional | Male | 2.632 | - | 3.405 |
| contributions | | (0.523) | - | (0.605) |
| | Female | 4.054 | - | 4.842 |
| | | (0.613) | - | (0.493) |
| (B) Expectation of other | Male | 3.35 | - | 3.21 |
| unconditional conts. | | (0.326) | - | (0.294) |
| | Female | 3.44 | - | 3.75 |
| | | (0.363) | - | (0.320) |
| (C) Belief-normalized | Male | -0.718 | <* | 0.197 |
| unconditional conts. | | (0.422) | - | (0.458) |
| | Female | 0.610 | - | 1.090 |
| | | (0.385) | >* | (0.294) |
| (D) Conditional | Male | 2.659 | - | 3.192 |
| contributions | | (0.478) | - | (0.509) |
| (aggregated to | Female | 3.369 | - | 4.078 |
| individual level) | | (0.598) | >* | (0.471) |

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Recall our theoretical predictions. Theorem 1 suggests that females will react to public observation by moving toward actions that signal an average type, while males will choose actions that signal higher types. This leads to two predictions, namely, visibility will: increase average male giving (Proposition 1) and decrease the variance among female giving (Proposition 2). To address these predictions, we report the impact of visibility on both the means and variances of our outcome variables in Table 4. In this table, differences in means are tested using one-sided t-tests (allowing for unequal variances). Differences in variance are tested using one-sided variance-ratio tests.

Comparing the average and variances of unconditional contribution ($c_i^u$) across treatments in Panel A reveals minimal treatment effects. Average contributions are slightly higher for both genders in the *Visibility* treatment. Variance appears slightly higher among male contributions and slightly lower among female contributions. None of these differences are statistically significant at the 10% level. However, this comparison does not take into account heterogeneity in subjects' perception of norms (Panel B). Without any information about the audience or hypothetical scenarios to respond to, a subject in the unconditional contribution phase can only base her reaction to the visibility treatment on her own beliefs ($E_i[c_{-i}^u]$). The belief-normalized contributions ($c_i^u - E_i[c_{-i}^u]$) in Panel C correct for the assumption of homogenous beliefs in Panel A. The quantities can be interpreted as participants' reaction to what they expect others to give. Figure 5 displays the distribution of belief-normalized contributions.

**FIGURE 5**
**Distribution of belief-normalized contributions**



We find that on average, males give 70 cents less than what they expect others to give in the control treatment and 20 cents more in the visibility treatment. Females give 60 cents more than what they expect others to give in the control treatment and $1 more in the visibility treatment. Consistent with Proposition 1, we find that the increase in contribution is significant for males and not for females. We confirm the treatment effect in a set of regression of the form

$$c_i^u = \alpha + \beta_v(visibility_i) + \beta_b[beliefs_i] + \beta_f(unfamiliar_i) \quad (8)$$

where $[beliefs]$ is the vector of elicited probabilities and $unfamiliar$ is a dummy variable that indicates that the participant is unfamiliar with the charitable cause (and is hence less sympathetic to it). Confirming the results in Panel C, we see that the coefficient $\beta_v$ is positive and significant for males (Column (1)), but not for females (Column (2)). The coefficients on beliefs and familiarity are highly significant, indicating that these factors are important determinants of contribution.

**Lab result 1:** Male subjects respond to visibility by increasing their unconditional contribution with respect to their beliefs about others' contribution.

We now turn our attention to Proposition 2. The distribution of belief-normalized contributions in Figure 5 illustrates that females' unconditional contributions in the visibility treatment are concentrated near what they expect others to give. This suggests a decrease in variance in unconditional contributions, which is confirmed in the variance-ratio test in Panel C.[17] Visibility does not affect variance with male contributions.

**Lab result 2:** Visibility decreases the variance of belief-normalized unconditional contributions among female subjects.

---

[17] We obtain similar results when we use other measures of central tendency (mode and median) to normalize beliefs. See Table A.2 in Appendix A.

**TABLE 5**
**OLS – Initial assessment of treatment effects**

| VARIABLES | (1)<br>Uncond. cont.<br>Males only | (2)<br>Uncond. Cont.<br>Females only | (3)<br>Cond. Cont.<br>Males only | (4)<br>Cond. Cont.<br>Females only |
|---|---|---|---|---|
| Visibility | 1.309** | 0.600 | 0.913* | 0.521 |
| | (0.612) | (0.503) | (0.470) | (0.476) |
| Prob(0) | -11.02*** | -13.05*** | -11.45*** | -12.32*** |
| | (2.587) | (2.773) | (2.448) | (2.888) |
| Prob(2) | -11.71*** | -13.62*** | -12.68*** | -12.22*** |
| | (2.496) | (2.856) | (2.530) | (3.047) |
| Prob(4) | -10.46*** | -7.748** | -8.131** | -6.212* |
| | (3.162) | (3.087) | (3.092) | (3.347) |
| Prob(6) | -7.566** | -6.183* | -11.34*** | -5.237* |
| | (3.764) | (3.277) | (3.056) | (2.652) |
| Prob(8) | -0.589 | -4.763 | -1.990 | -4.790 |
| | (6.212) | (6.380) | (5.061) | (7.763) |
| Unfamiliar | -1.368** | 0.233 | -1.045** | -0.588 |
| | (0.630) | (0.513) | (0.498) | (0.508) |
| Constant | 12.62*** | 13.43*** | 12.91*** | 12.29*** |
| | (2.149) | (2.589) | (2.198) | (2.877) |
| Decision FE's | | | X | X |
| Observations | 75 | 75 | 1,575 | 1,575 |
| R-squared | 0.503 | 0.680 | 0.473 | 0.636 |

Standard errors in parentheses
(Standard errors clustered at individual-level in columns 3 and 4)
*** p<0.01, ** p<0.05, * p<0.1

We now test Propositions 1 and 2 in the context of conditional contribution ($c_i^c$). As in Panel A, Panel D suggests no statistically significant change in average conditional contribution for both genders. Visibility has no effect on male contribution variance but significantly decreases female contribution variance. However, this simple comparison of means may not capture the true treatment effects since it does not take into account the fact that subjects are making a series of 21 decisions. It also does not account for individual beliefs, which will be important if subjects consider some scenarios to be implausible. For instance, a subject may think that it is highly unlikely that both group members gave $10 and hence may not respond to ($10, $10) in the same way she responds to ($4, $2).

In Columns 3 and 4 of Table 5, we use a regression to assess the impact of visibility on size of conditional contributions while accounting for beliefs and cause familiarity. Let $c_{id}^C$ indicate participant $i$ contribution in conditional decision $d \in \{1,..,21\}$ and $\delta_d$ be a set of decision fixed effects. We estimate the equation below:

$$c_{id}^C = \alpha + \beta_v(visibility_i) + \delta_d + \beta_b[beliefs_i] + \beta_f(unfamiliar_i) \quad (9)$$

Paralleling the unconditional results, we find that visibility increases male giving (3) but not female giving (4). We also find that elicited beliefs are indeed important predictors of conditional contributions. We use a similar regression to confirm the decrease in variance among female conditional contributions in Table A.3 in Appendix A.[18]

**Lab result 3:** Visibility increases male conditional contributions and decreases the variance of conditional contributions among female subjects.

### 4.2.2 Females Choice of Middle Action (Corollary to Theorem 1)

A decrease in variance among female contributions shows that visibility induces females to choose a particular action. However, this decrease alone does not necessarily show that females have wallflower reputational concerns, since it is possible that female contributions are converging on choices other than those that signal average type.

In this section we use conditional contributions to test if the visibility treatment induces females to choose the middle action, as is implied in Corollary (i). Let the hypothetical contributions (from an individual's two other group members) be denoted as $c_{min}$ and $c_{max}$, with $c_{min} \leq c_{max}$. If females respond to visibility by avoiding stigma and honor while males seek honor and avoid stigma, then the following relationships should hold:

$$\Pr(c_{min} < u^c < c_{max} \mid F, vis) > \Pr(c_{min} < u^c < c_{max} \mid F, control) \quad (10)$$

[18] In Table A.3, we regress the absolute deviation of conditional contribution from the mean group contribution $|c_i^c - (c_1^u + c_2^u)/2|$ against the set of control in Table 5. We find that visibility reduces variance among female contributions, but not among male contributions.

$$\Pr(c_{min} < u^c < c_{max} \mid F, vis) > \Pr(c_{min} < u^c < c_{max} \mid M, vis)$$

To test Eq (10), we estimate a series of five linear probability models assessing the likelihood that, for some conditional decision, a participant chooses a contribution within a particular range relative to her group members' (hypothetical) contributions. These ranges are: (1) less than the minimum of group members' contributions, (2) equal to the minimum, (3) within the minimum and maximum, (4) equal to the maximum, and (5) greater than the maximum. Thus, we estimate a set of regressions of the form:

$$y_{id}^r = \alpha + \beta_f(female_i) + \beta_v(visibility_i) + \beta_{fv}(female_i \times visibility_i) + \delta_d + \beta_b[beliefs_i] + \beta_f unfamiliar_i$$

where $y_{id}^r = 1$ if participant $i$ chooses a contribution within the associated range $r$ in conditional decision $d \in \{1, .., 21\}$. Note that it is not always possible to choose a contribution within a particular range. For instance, when the minimum of partners' contributions is 0, it is impossible to choose a contribution "less than the minimum." These instances are excluded from the analysis.[19] As before we include $\delta_d$ (decision fixed effects) to control for beliefs and familiarity with the cause and we cluster standard errors at the individual level.

The coefficient of primary interest is $\beta_{fv}$, as it captures the *gender difference in response to visibility*. Also of interest is the sum of the *Visibility* and *Female X Visibility* coefficients ($\beta_v + \beta_{fv}$), which is the treatment effect conditional on being female. (The treatment effect conditional on being male is the visibility coefficient.) Our hypothesis suggests that visible females should be more likely to choose a contribution within the range of their partners' contributions to avoid both stigma and honor. Thus, we would expect $\beta_{fv}$ to be positive in column (3).

[19] Other exclusions: It is impossible to be "greater than maximum" when the maximum contribution is 10, so all such cases are excluded from "greater than max." regression. It is impossible to be strictly "within range" when the distance between partners' contributions is 0 or 2, so these cases are excluded from "within range" regression. Finally, there is no unique maximum or minimum when partners' contributions are identical, so these cases are excluded from the "equal to min." and "equal to max." regressions.

**TABLE 6**
**Likelihood of choosing a conditional contribution within a particular range**

| VARIABLES | (1)<br><min | (2)<br>=min | (3)<br>In range | (4)<br>=max | (5)<br>> max |
|---|---|---|---|---|---|
| Female X Visibility | -0.00547 | -0.0157 | 0.160** | -0.0381 | -0.0170 |
| | (0.0811) | (0.0372) | (0.0731) | (0.0406) | (0.0701) |
| Female | -0.0559 | -0.0133 | -0.0533 | 0.0524* | 0.0236 |
| | (0.0604) | (0.0214) | (0.0495) | (0.0295) | (0.0486) |
| Visibility | -0.114* | 0.0259 | -0.0373 | 0.0505** | 0.0230 |
| | (0.0593) | (0.0288) | (0.0547) | (0.0219) | (0.0483) |
| | *Controls for cause familiarity, the vector of elicited probabilities, and decision fixed effects are included but not displayed.* | | | | |
| Constant | -0.166 | -0.107* | -0.158 | 0.425*** | 1.334*** |
| | (0.170) | (0.0621) | (0.133) | (0.0952) | (0.187) |
| *Female X Vis. + Vis.* | -0.119** | 0.010 | 0.122** | 0.012 | 0.006 |
| | (0.053) | (0.025) | (0.051) | (0.035) | (0.051) |
| Observations | 2,250 | 2,250 | 1,500 | 2,250 | 2,250 |
| R-squared | 0.452 | 0.098 | 0.232 | 0.128 | 0.429 |

Robust standard errors (clustered at individual-level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Indeed in we do find *Female X Visibility* to be positive and significant in Table 6 column (3), confirming that female subjects are choosing the “middle action” in the visibility treatment. Interestingly, *Female* is positive and significant in column (4), suggesting that females were comfortable contributing at the maximum of the group range in private. This pattern of contributing above the norm in private and conforming to the norm in public is consistent with our field observations where we saw females completing multiple word search puzzles in private and completing exactly one puzzle when their identities would be revealed.

Visibility has a different effect on males. While revelation of identity induces both genders to avoid stigma, column (4) in Table 6 indicates that it induces only males to choose a contribution that matches the maximum of their group members’ contributions.[20] Altogether, these findings confirm the existence of gender differences in reputation concerns in response to public observation.

[20] Coefficients for Visibility and *Female X Vis. + Vis.* in Column (1) indicate that both genders avoid contributing below the range in the visibility treatment.

**Lab result 4:** Visibility induces females to choose the "middle" action and males to choose the high action.

## 5. Conclusion

In this paper, we address the question of how public recognition impacts prosocial behavior. We follow a wealth of theoretical and experimental research on the topic. However, despite the large existing literature, there is much that remains unclear. Existing research has found that in some instances, visibility clearly increases prosocial behavior; in other cases, it clearly decreases prosocial behavior. In still other situations the impact of visibility is dependent upon additional details of the environment. It is with this in mind that we focus less on the question of whether visibility increases or decreases prosocial behavior and more on the mechanisms through which visibility acts. We focus in particular on the roles (and interaction) of conformity and reputation concerns.

In our field experiment, we find that females and males react differently to visibility in a prosocial environment. Males contribute more to a charity when their contributions are visible, while females conform to an established norm. Building on Benabou & Tirole's 2006 model of prosocial behavior, we suggest that this result may be driven by a difference in males' and females' preferences over reputation. While males may be more likely to have the "classic" Benabou & Tirole motivation to avoid stigma and seek honor, females are more likely to prefer to avoid both stigma and honor. That is, females do not want to appear as though they are *not* prosocial by choosing a low contribution, but they also prefer that their contribution not be perceived as being strictly motivated by reputation or prestige. As such, they choose an action that draws minimal attention by conforming to a norm that minimizes both of these forms of reputation.

Thus, our more general claim in this paper is that the impact of visibility – and whether it increases or decreases prosocial behavior – depends heavily on the location of the norm. In settings where there is a norm of high giving with visibility (as in public goods games), visibility pulls the contributions of both honor-seeking and wallflower types upwards; honor-seeking types choose a high contribution to signal altruism while wallflowers choose a high contribution to conform. However, where there is a separation between the "most generous" action and the norm (as in a dictator game, where there is typically a norm of equal division (Andreoni and Bernheim, 2009)), we might expect separation between the observed behavior of honor-seeking types and wallflower types. This would diminish (or even reverse) the benefits of visibility.

With this explanation in mind, we conduct a laboratory experiment that allows us to observe participants under a variety of norms. We do this in order to draw a causal link between the location of the norm and the directional impact of visibility on prosocial behavior. In the lab, participants must decide how much money to give to a charitable cause, but are given the opportunity to condition their contribution on the contributions of their randomly assigned group members, who – in one treatment – will eventually observe their contribution. We find that females react to visibility by choosing a contribution that is *neither above nor below* the contributions of her group members. This is consistent with the idea that females seek to avoid appearing either selfish or reputation-motivated. Additionally, we find that visibility also reduces the dispersion of females' contributions around what they expect others to do.

These findings are far from being of purely theoretical interest. Outside of the experimental laboratory, prosocial behavior is rarely anonymous. Fundraisers often purposefully increase the visibility of contributions by publishing donor lists for instance. Other activities – such as volunteering or going to a polling location to vote, which are more often than not carried out in public and in the presence of peers – are visible by their very nature. Thus, a deeper understanding of the impact of such visibility is critical.

Our findings suggest that care must be taken in manipulating the degree of visibility of prosocial behavior. This is particularly true in prosocial environments where females are overrepresented, as is the case in the nonprofit sector. Indeed, as Gibelman (2000) demonstrates, though females make up a large majority of the nonprofit sector, upper management positions continue to be male-dominated. While we certainly do not claim to fully explain this phenomenon, our finding that females prefer to avoid drawing attention to themselves when prosocial behavior is visible, may play an important role. In the context of charitable giving, males and females have different preferences in the charitable causes they support (Andreoni et al., 2003). This fact, combined with our results, suggests that causes that attract more female givers may require a very different fundraising strategy than those that attract males.

**Acknowledgement**

We thank Lise Vesterlund for very helpful discussions. We also thank participants at Pittsburgh Experimental Brownbag, at the ESA 2011 meeting (Tuscon), the 2011 ARNOVA meeting (Toronto), the 2012 Workshop on Social Dimension of Organizations, and seminar participants at Universitat Bonn for helpful comments. We thank GSPIA for their financial support (SL) and Kate Morris for her careful reading of the draft.

## References

F. Alpizar, F. Carlsson, and O. Johansson-Stenman, 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics*, 92(5-6):1047--1060.

J. Andreoni and B.D. Bernheim, 2009. Social image and the 50--50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607--1636.

J. Andreoni, E. Brown, and I. Rischall, 2003. Charitable giving by married couples: who decides and why does it matter?. *Journal of Human Resources*, 38(1):111.

J. Andreoni and R. Petrie, 2004. Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics*, 88(7-8):1605--1623.

D. Ariely, A. Bracha, and S. Meier, 2009. Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, 99(1):544--555.

C. Bicchieri, 2006. The grammar of society. *Cambridge Books*.

R. Bénabou and J. Tirole, 2006. Incentives and prosocial behavior. *The American economic review*, 96(5):1652--1678.

R.B. Cialdini, R.R. Reno, and C.A. Kallgren, 1990. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places.. *Journal of personality and social psychology*, 58(6):1015.

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

R. Croson, F. Handy, and J. Shang, 2009. Keeping up with the joneses: the relationship of perceived descriptive social norms, social information, and charitable giving. *Nonprofit Management and Leadership*, 19(4):467--489.

M. Dufwenberg and A. Muren, 2006. Generosity, anonymity, gender. *Journal of Economic Behavior & Organization*, 61(1):42--49.

U. Fischbacher, S. Gachter, and E. Fehr, 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397--404.

S. Gachter and E. Fehr, 1999. Collective action as a social exchange. *Journal of Economic Behavior & Organization*, 39(4):341--369.

M. Gibelman, 2000. The Nonprofit Sector and Gender Discrimination. *Nonprofit Management and Leadership*, 10(3):251--269.

U. Gneezy, M. Niederle, and A. Rustichini, 2003. Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, 118(3):1049--1074.

U. Gneezy and A. Rustichini, 2004. Gender and competition at a young age. *American Economic Review*, 94(2):377--381.

H. Hall, 2004. Gender differences in giving: going, going, gone?. *New Directions for Philanthropic Fundraising*, 2004(43):71--81.

K. Kanthak and Jonathan Woon. 2012. "Women Don't Run? Gender Differences in Candidate Entry" Working Paper: University of Pittsburgh.

S. Linardi and Margaret Anne McConnell. 2008, "Volunteering and image concerns", California Institute of Technology Social Science Working Paper #1282, Division of the Humanities and Social Sciences.

S. Linardi and M.A. McConnell, 2011. No excuses for good behavior: volunteering and the social environment. *Journal of Public Economics*, 95(5-6):445--454.

C. Mellstrom and M. Johannesson, 2008. Crowding out in blood donation: was Titmuss right?. *Journal of the European Economic Association*, 6(4):845--863.

C. Noussair and S. Tucker, 2007. Public observability of decisions and voluntary contributions in a multiperiod context. *Public Finance Review*, 35(2):176--198.

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

X.S. Pan and D. Houser, 2011. Competition for trophies triggers male generosity. *PloS one*, 6(4):e18050.

M. Rege and K. Telle, 2004. The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, 88(7-8):1625--1644.

L. Shi, 2011. Monetary rewards, image concern, and intrinsic motivation: evidence from a survey on blood donation. *Working Papers*.

A.R. Soetevent, 2005. Anonymity in giving in a natural context--a field experiment in 30 churches. *Journal of Public Economics*, 89(11-12):2301--2323.

M.A. Taylor and A.R. Kaminski, 1997. Women as philanthropists: leading the transformation of fundraising. *International Journal of Nonprofit and Voluntary Sector Marketing*, 2(2):107--113.

D. Zetland and M. Della Giusta, 2011. Focal points, gender norms and reciprocation in public good games. *Economics & Management Discussion Papers*.

**APPENDIX A: Additional figures and screenshots**

**FIGURE A.1**
**Initial Screen**

**FIGURE A.2**
**Scores treatment – observation of previous donations**

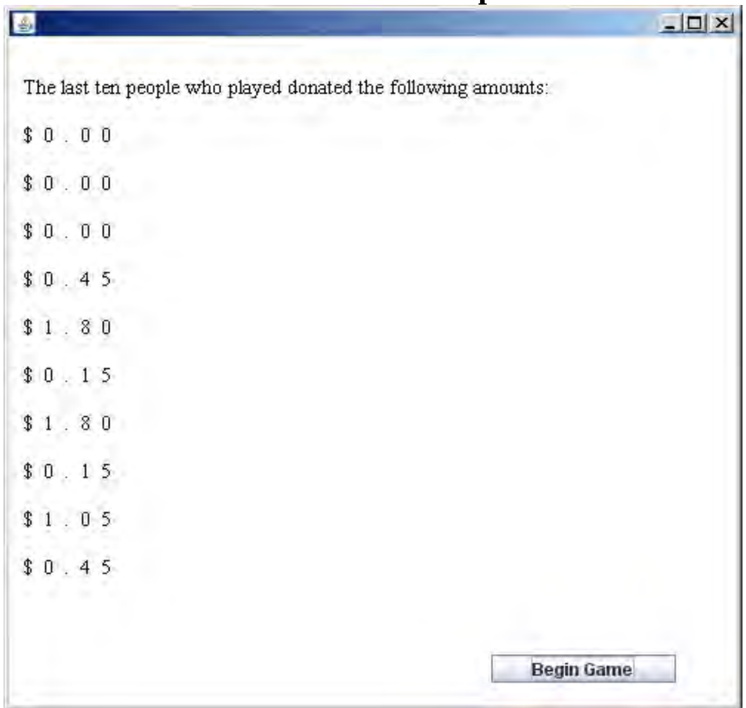



**FIGURE A.3**
**Names treatment – name entry**

**FIGURE A.4**
**Names treatment – observation of previous donations**


**FIGURE A.5**
**Post-experiment questionnaire**

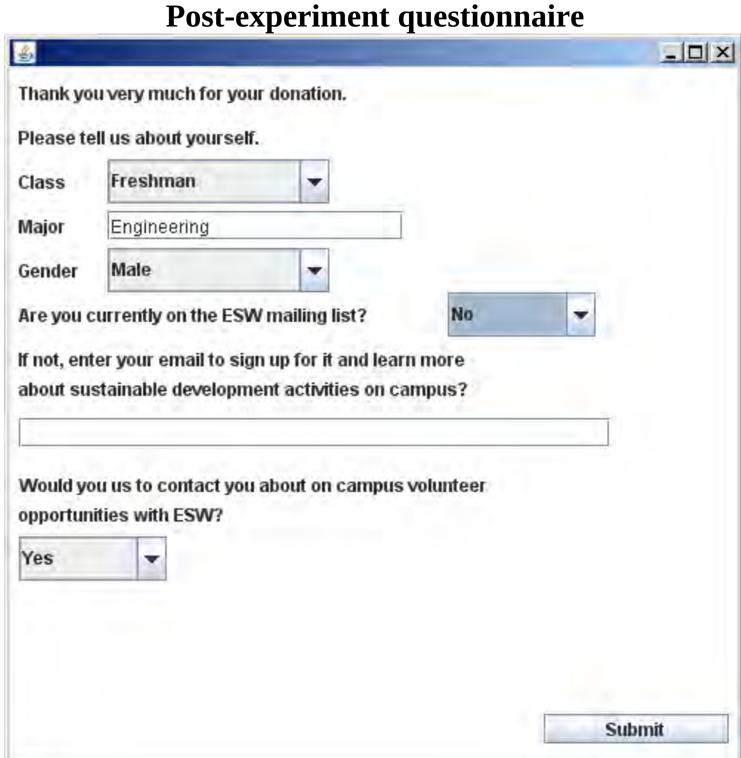


**FIGURE A.6**
**Distribution of donations**

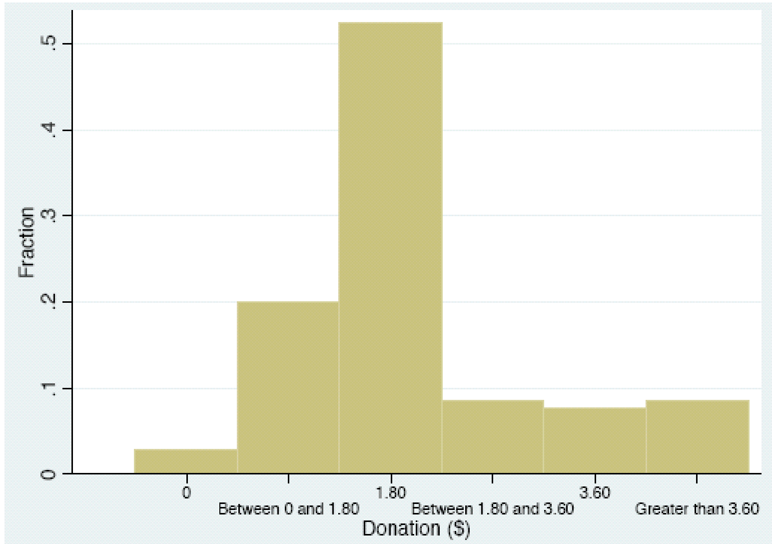



**FIGURE A.7**
**Mean of donations of by treatment and gender**

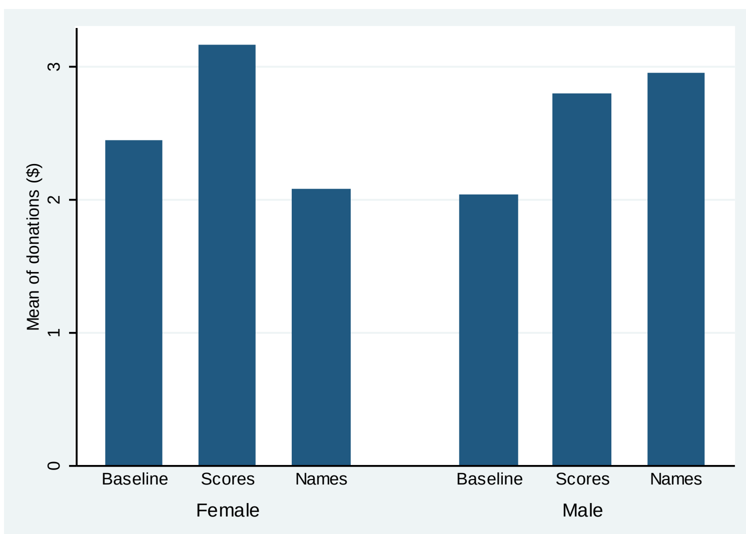


**TABLE A.1**
**Summary of survey responses**

| | Freq. | Percent | | Freq. | Percent |
|---|---|---|---|---|---|
| **Gender** | | | **Class** | | |
| Female | 29 | 27.88 | Freshman | 21 | 20.19 |
| Male | 46 | 44.23 | Sophomore | 4 | 3.85 |
| No response | 29 | 27.88 | Junior | 5 | 4.81 |
| | | | Senior | 12 | 11.54 |
| **Interest in volunteer opportunities** | | | Grad | 22 | 21.15 |
| No | 21 | 20.19 | Other | 6 | 5.77 |
| Yes | 21 | 20.19 | No response | 34 | 32.69 |
| No response | 62 | 59.62 | | | |

## TABLE A.2
## Mean & variance comparisons of belief-normalized unconditional contributions

| Belief-normalization: | | Control | Sig. diff.[1] | Visibility |
|---|---|---|---|---|
| $c^u_i$-$E_i(c^u_{-i})$ | Male | -0.718 | <* | 0.197 |
| | | (0.422) | - | (0.458) |
| | Female | 0.610 | - | 1.090 |
| | | (0.385) | >* | (0.294) |
| $c^u_i$-Median | Male | -0.500 | <** | 0.487 |
| | | (0.436) | - | (0.429) |
| | Female | 0.919 | - | 1.132 |
| | | (0.372) | - | (0.314) |
| $c^u_i$-Mode | Male | -0.158 | <* | 0.703 |
| | | (0.402) | - | (0.433) |
| | Female | 0.973 | - | 1.11 |
| | | (0.434) | >*** | (0.279) |

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
[1] *Tests of significant differences: Differences in means are tested using one-sided t-tests (allowing for unequal variances). Differences in variance are testing using one-sided variance-ratio tests.*

**TABLE A.3**
**Conditional contribution:**
**Absolute deviation from mean of partners' hypothetical contributions**

| | (1) | (2) |
|---|---|---|
| VARIABLES | Males only | Females only |
| | | |
| Visibility | -0.369 | -0.495** |
| | (0.260) | (0.239) |
| Prob(0) | -1.063 | -1.804* |
| | (0.741) | (1.080) |
| Prob(2) | -1.561* | -3.140*** |
| | (0.829) | (1.108) |
| Prob(4) | -2.823** | -5.694*** |
| | (1.300) | (1.515) |
| Prob(6) | -2.136 | -4.293*** |
| | (1.355) | (1.149) |
| Prob(8) | -2.750 | -5.312* |
| | (1.969) | (3.131) |
| Unfamiliar | 0.391 | 0.0111 |
| | (0.267) | (0.278) |
| Constant | 5.414*** | 7.216*** |
| | (0.556) | (1.019) |
| | | |
| Decision FE's | X | X |
| | | |
| Observations | 1,575 | 1,575 |
| R-squared | 0.473 | 0.636 |

Robust standard errors (clustered at individual-level) in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**Appendix B: "Wallflower" model proofs**

Individuals choose between contribution level $a \in \{0,1,2\}$ where $c(a) = ka^2$. Their utility function is:

$$u(a|g,x,v) = va - c(a) + xR(a|g) \tag{B.1}$$

where

$$\begin{aligned} &R(a|g) = h(g)max[E(v|a,g) - \bar{v}, 0] - max[\bar{v} - E(v|a,g), 0] \\ &h(g) = 1 \textit{ when } g = M \textit{ and } h(g) = -1 \textit{ when } g = F \end{aligned} \tag{B.2}$$

With uniform distribution, the expected type of an individual who gives $a$ is the midpoint of the cutoff types for $a$ and $a+1$:

$$\begin{aligned} E(v|0,g,x) &= \frac{\tilde{v}_1^g}{2} \\ E(v|1,g,x) &= \frac{\tilde{v}_1^g + \tilde{v}_2^g}{2} \\ E(v|2,g,x) &= \frac{\tilde{v}_2^g + A}{2} \end{aligned} \tag{B.3}$$

Since $E(v|0,g,x) \leq \bar{v}$ and $E(v|2,g,x) \geq \bar{v}$, we can write $R(0|g)$ as $\frac{\tilde{v}_1^g - A}{2}$ and $R(2|g)$ as $\frac{\tilde{v}_2^g}{2}$ for male and $-\frac{\tilde{v}_2^g}{2}$ for females.

Define marginal reputation benefit as $r(a|g) \equiv R(a|g) - R(a-1|g)$. For males, $R(1|M) = \frac{\tilde{v}_1^M + \tilde{v}_2^M - A}{2}$ whether $\tilde{v}_1^M$ is less than or larger than $\bar{v}$. Male marginal reputation for contributing a higher amount is always positive.

$$\begin{aligned} r(1|M) &= \frac{\tilde{v}_2^M}{2} > 0 \\ r(2|M) &= \frac{A - \tilde{v}_1^M}{2} > 0 \end{aligned} \tag{B.4}$$

For females $R(1|F)$ depends on whether $E(v|1,F,x)$ is above and below $\bar{v}$, the average type. Consider the first case. $E(v|1,F,x) > \bar{v}$ implies $R(1|F) = \frac{A - \tilde{v}_1^F - \tilde{v}_2^F}{2}$. Here, since $a = 1$ already signals a type above the average type, the marginal reputation benefit of increasing contribution to $a = 2$ only further intensifies (unwanted) image signals ($r(2|F) < 0$).

$$\begin{aligned} r(1|F) &= A - \tilde{v}_1^F - \frac{\tilde{v}_2^F}{2} \\ r(2|F) &= \frac{\tilde{v}_1^F - A}{2} < 0 \end{aligned} \tag{B.5}$$

Now consider the latter case. When $E(v|1,F,x) < \bar{v}$, $R(1|F) = \frac{\tilde{v}_1^F + \tilde{v}_2^F - A}{2}$. Here, since $a = 1$ already signals a type below the average type, decreasing contribution from $a = 1$ to $a = 0$ will further intensifies stigma. For those contributing $a = 0$, the marginal reputation benefit of increasing contribution to $a = 1$ is therefore positive ($r(1|F) > 0$).

$$r(1|F) = \frac{\tilde{v}_2^F}{2} > 0$$
$$r(2|F) = \frac{A - \tilde{v}_1^F}{2} - \tilde{v}_2^F \tag{B.6}$$

**Lemma 1** $A \geq c_2 \Rightarrow E(v|a=1, x=1) \leq \bar{v}$

*Proof.* Suppose $A \geq c_2$ and $E(v|a=1) > \bar{v}$. Then using the definition of cutoff types: $v_a^g = c(a) - c(a-1) - xr(a|g)$ and Eq. B.5, we solve for female cutoff types for visible contributions:

$$\tilde{v}_1^F = c_1 - r(1|F) = c_1 - A + \tilde{v}_1^F + \frac{\tilde{v}_2^F}{2}$$

$$\tilde{v}_2^F = c_2 - c_1 - r(2|F) = c_2 - c_1 - \frac{\tilde{v}_1^F - A}{2}$$

The first equation implies

$$\tilde{v}_2^F = 2(A - c_1) \tag{B.7}$$

Substituting this to the second equation above, we arrive at

$$\tilde{v}_1^F = 2(c_1 + c_2 - \frac{3}{2}A) \tag{B.8}$$

Substituting $\tilde{v}_1^F$ and $\tilde{v}_2^F$ into our assumption that $E(v|a=1) = \frac{\tilde{v}_1^g + \tilde{v}_2^g}{2} > \bar{v} = \frac{A}{2}$, we arrive at

$$c_1 + c_2 - \frac{3}{2}A + A - c_1 > \frac{A}{2}$$

The above equation implies $A < c_2$ which is a contradiction. □

**Lemma 2** When contributions are visible, the cutoffs types are:

$$\tilde{v}_1^M = \frac{2}{5}(\frac{A}{2} - c_2 + 3c_1) \qquad \tilde{v}_1^F = A - 2c_2 + 2c_1$$
$$\tilde{v}_2^M = \frac{2}{5}(2c_2 - c_1 - A) \qquad \tilde{v}_2^F = 2(2c_2 - c_1 - A) \tag{B.9}$$

*Proof.* When contributions are visible $\tilde{v}_2^g = c(a) - c(a-1) - r(a|g)$. By Lemma 1 we know that $A \geq c_2 \Rightarrow E(v|a=1, x=1) \leq \bar{v}$, which means we only have to be concerned with female reputation as defined by Eq. B.6. Since $r(1|g) = \frac{\tilde{v}_2^g}{2}$ for both gender:

$$\tilde{v}_1^g = c_1 - r(1|g) = c_1 - \frac{\tilde{v}_2^g}{2}$$
$$\tilde{v}_2^g = 2(c_1 - \tilde{v}_1^g) \tag{B.10}$$

However, $r(2|g)$ is gender specific. For male this is:

$$\tilde{v}_2^M = c_2 - c_1 - r(2|M) = c_2 - c_1 - \frac{A - \tilde{v}_1^M}{2} \tag{B.11}$$

Setting Eq. B.10 equal to Eq. B.11, we get $\tilde{v}_1^M = \frac{2}{5}(\frac{A}{2} - c_2 + 3c_1)$, which we substitute back to Eq. B.10 to arrive at $\tilde{v}_2^M$.
For females:

$$\tilde{v}_2^F = c_2 - c_1 - r(2|F) = c_2 - c_1 - \frac{A - \tilde{v}_1^F}{2} + \tilde{v}_2^F \tag{B.12}$$

From Eq. B.12 we arrive at $\tilde{v}_1^F = A + 2c_1 - 2c_2$ which we substitute to Eq. B.10 to arrive at $\tilde{v}_2^F$. □

We first rewrite Theorem 1 from the main text into three parts.

**Theorem 1**

(i) When contributions are not visible, $v_1 = v_1^M = v_1^F = c_1$, and $v_2 = v_2^M = v_2^F = c_2 - c_1$

(ii) When contributions are visible, the cutoffs are well behaved when A is not to large relative to costs of contribution:

$$\tilde{v}_1^M < \tilde{v}_2^M \textit{ and } \tilde{v}_1^F < \tilde{v}_2^F \Leftrightarrow A < 2c_2 - \frac{8}{3}c_1 \tag{B.13}$$

(iii) Within this range, $\tilde{v}_1^F < \tilde{v}_1^M < v_1 < \tilde{v}_2^M < \bar{v} < v_2 < \tilde{v}_2^F$

*Proof.* (i) Since $xr(a|g) = 0$ for when contributions are not visible, $v_a^g = c(a) - c(a-1)$. This means $v_1^M = v_1^F = c_1$, and $v_2^M = v_2^F = c_2 - c_1$.

(ii) Cutoff types are well behaved when $\tilde{v}_1^g < \tilde{v}_2^g$. Substituting cutoffs for males from Lemma 2:

$$\tilde{v}_1^M = \frac{2}{5}(\frac{A}{2} - c_2 + 3c_1) < \tilde{v}_2^M = \frac{2}{5}(2c_2 - c_1 - A)$$

$$\frac{A}{2} - c_2 + 3c_1 < 2c_2 - c_1 - A$$

Hence we arrive at the following condition:

$$A < 2c_2 - \frac{8}{3}c_1 \Rightarrow \tilde{v}_1^M < \tilde{v}_2^M \tag{B.14}$$

Substituting cutoffs for females from Lemma 2 we see that the condition where cutoff types are well behaved for female contributors is satisfied automatically when Eq. B.14 is satisfied:

$$\tilde{v}_1^F = A - 2c_2 + 2c_1 < \tilde{v}_2^F = 2(2c_2 - c_1 - A)$$

$$3A < 2(2c_2 - c_1) + 2c_2 - 2c_1$$

$$A < 2c_2 - \frac{4}{3}c_1$$

(iii) We restrict our attention to $c_2 < A < 2c_2 - \frac{8}{3}c_1$ where cutoffs are well behaved. First note that $v_1 < \bar{v} = \frac{A}{2}$ since $c_1 < \frac{c_2}{2}$ by assumption and that $\bar{v} < v_2$ since $2c_2 - \frac{8}{3}c_1 < 2(c_2 - c_1)$.
We now compare the cutoff type for $a = 2$. Lemma 2 directly shows that $\tilde{v}_2^M < \tilde{v}_2^F$ so we will show that $\tilde{v}_2^M < \frac{A}{2}$ and $v_2 < \tilde{v}_2^F$. Starting from the first inequality:

$$\tilde{v}_2^M = \frac{2}{5}(2c_2 - c_1 - A) < \frac{c_2}{2} < \frac{A}{2}$$

$$2c_2 - c_1 - A < \frac{5c_2}{4}$$

$$\frac{3}{4}c_2 - c_1 < A$$

which is true since $A > c_2$. Note also that $\tilde{v}_2^M > 0$ since $A < 2c_2 - \frac{8}{3}c_1$. Now turning to the latter, since $\tilde{v}_2^F$ decrease in $A$, we substitute the upper bound of $A$ to the inequality below:

$$c_2 - c_1 = v_2 < \tilde{v}_2^F = 2(2c_2 - c_1 - A)$$

$$\frac{c_2}{2} - \frac{c_1}{2} < 2c_2 - c_1 - 2c_2 + \frac{8}{3}c_1$$

$$c_2 < \frac{19}{3}c_1$$

which is true for since $c_2 = 4c_1$
Lastly we compare cutoffs for $a = 1$. Since $\tilde{v}_2^M > 0$, by Eq. B.10, $\tilde{v}_1^M > v_1^M$. We only need to show that $\tilde{v}_1^F < \tilde{v}_1^M$

$$\tilde{v}_1^F = A - 2c_2 + 2c_1 < \tilde{v}_1^M = \frac{2}{5}(\frac{A}{2} - c_2 + 3c_1)$$

$$\frac{4}{5}A < \frac{2}{5}(-c_2 + 3c_1) + 2c_2 - 2c_1 = \frac{4}{5}(2c_2 - c_1)$$

which is true when Eq. B.14 is satisfied.

□

**Proposition 1**

*Proof.* Let $\bar{a}^g$ be the average contribution of gender g .

$$\bar{a}^g = 1\frac{v_2^g - v_1^g}{A} + 2\frac{A - v_2^g}{A}$$

Simplifying we arrive at:

$$\tilde{\bar{a}}^g = 2A - \frac{v_2^g + v_1^g}{A}$$

which is decreasing in the sum of the two cutoffs $v_2^g$ and $v_1^g$. When contributions are not visible, by Theorem 1(i) $v_2^g + v_1^g = c_2$. For male, when contributions are visible:

$$\tilde{v}_2^M + \tilde{v}_1^M = \frac{2}{5}(c_2 - 2c_1 - \frac{A}{2}) << c_2$$

For female, substituting the bounds for A into the sum of cutoffs when contributions are visible indicate that the impact of visibility on contributions is less positive:

$$\frac{8}{3}c_1 < \tilde{v}_2^g + \tilde{v}_1^g = 2c_2 - A \leq c_2$$

□

**Corollary to Theorem 1** Let $a_g^*(\bar{v})$ and $\widetilde{a}_g^*(\bar{v})$ be the actions chosen by the median (average) type of gender $g$ when actions are not visible and visible, respectively.

$$a_F^*(\bar{v}) = \widetilde{a}_F^*(\bar{v}) = a_M^*(\bar{v}) = 1 < \widetilde{a}_M^*(\bar{v}) = 2$$

**Proposition 2**

*Proof.* The median type's expected contribution is $a_g^*(\bar{v})$ and $\widetilde{a}_g^*(\bar{v})$ for non-visible and visible contribution, respectively. Average squared distance between an individual's contributions and the median type's expected contribution

$$(0-\bar{a}_g)^2 v_1^g + (1-\bar{a}^g)^2(v_2^g - v_1^g) + (2-\bar{a}^g)^2(A - v_2^g)$$

By Corollary to Theorem 1, $\widetilde{a}_M^*(\bar{v}) = 2$ implies the squared distance of $a = 0$ and $a = 1$ to average contribution have increased. However, the squared distance of $a = 2$ have decreased and since $\tilde{v}_a^M < v_a^M$, most of the weight is on $a = 2$. The effect of visibility on the variance of male contribution is therefore uncertain.
For females, since $\widetilde{a}_F^*(\bar{v}) = a_F^*(\bar{v}) = 1$, the squared distance for all $a$ remains (roughly) the same. However, since middle term decreases variance and the outer terms increases variance, $\tilde{v}_2^F - \tilde{v}_1^F > v_2^F - v_1^F$ decreases the variance for female contribution. □

## Appendix C: Incentive compatibility of beliefs

In the belief elicitation task, participants guess how many people chose each particular contribution level. Let $g_c$ denote a participant's reported guess as to how many people chose contribution level $c \in \{1,...C\}$. Let $n_c$ denote the actual number of participants who chose that contribution level. Participants' guesses are constrained such that $0 \leq g_c \leq N$ and $\sum_c g_c = N$, where $N$ is the number of other participants in the room. Payoffs are given by $\pi(g) = 0.5 \sum_c min\{g_c, n_c\}$.

Note that because participants have no information about other participants' actions when they make their guesses, their beliefs over the set of all outcomes $\{n_1,...,n_C\}$ are represented by $N$ i.i.d. draws from a multinomial distribution. Let $\{p_1,...,p_C\}$ denote the probabilities of each independent event $c$ that define the multinomial distribution, with $0 \leq p_c \leq N$ and $\sum_c p_c = 1$.

Given these parameters, denote the expected outcome as

$$\{y_1,...,y_C\} = E[\{n_1,...,n_C\}|\{N; p_1,...,p_C\}] \tag{C.1}$$

Note that $\{y_1,...,y_C\} = E[\{n_1,...,n_C\}|\{N; p_1,...,p_C\}] \Rightarrow \{y_1,...,y_C\} = \{Np_1,...,Np_C\}$.

**Claim 1**

If $\{y_1,...,y_C\} = E[\{n_1,...,n_C\}|\{N; p_1,...,p_C\}]$ and $\{n_1,...,n_C\}$ is generated according to a multinomial distribution defined by parameters $\{N; p_1,...,p_C\}$, then

$$\{y_1,...,y_C\} = argmax_g E[0.5 \sum_c min\{g_c, n_c\}]$$

*Proof.* We will prove by contradiction. Suppose that $\{y_1,...,y_C\} \neq argmax_g E[0.5 \sum_c min\{g_c, n_c\}]$. This implies that there is some rearrangement of guesses that would increase the participant's expected payoff. More specifically, there exists at least one pair of contribution choices $c$ and $c'$ such that the vector of guesses $\hat{y} = \{y_1,...,y_c+1, y_{c'}-1,...,y_C\}$ yields a higher expected payoff than $\bar{y} = \{y_1,...,y_c, y_{c'},...,y_C\}$. Without loss of generality, we let $c = 1$ and $c' = 2$.

Note that $\{y_1,...,y_C\} = E[\{n_1,...,n_C\}|\{N; p_1,...,p_C\}] \Rightarrow \{y_1,...,y_C\} = \{Np_1,...,Np_C\}$.

Thus, $\bar{y} = \{Np_1, Np_2, Np_3,...,Np_C\}$ and $\hat{y} = \{Np_1+1, Np_2-1, Np_3,...,Np_C\}$.

The expected payoff $E[0.5 \sum_c min\{g_c, n_c\}]$ can be rewritten as:

$$\sum_c g_c(1 - p(n_c \leq g_c)) + E[n_c|n_c \leq g_c]p(n_c \leq g_c) = \sum_c g_c(1 - F(g_c)) + E[n_c|n_c \leq g_c]F(g_c) \tag{C.2}$$

The marginal distribution of a single component of a multinomial is simply a binomial distribution, so $F(g_c)$ is the binomial CDF defined by parameters $(N, p_c)$. Next, note that $E[n_c|n_c \leq y_c] * F(y_c) = [\sum_0^{y_c} \frac{x*f(x)}{F(y_c)}] * F(y_c) = \sum_0^{y_c} x * f(x)$. Thus, the expected payoff can be further simplified:

$$\sum_c [g_c(1 - F(g_c)) + \sum_0^{y_c} x * f(x)] \tag{C.3}$$

The assumption $E\pi(\hat{y}) > E\pi(\bar{y})$ implies

$$(Np_1+1)(1-F(Np_1+1)) + \sum_{0}^{Np_1+1} x*f(x) + (Np_2-1)(1-F(Np_2-1)) + \sum_{0}^{Np_2-1} x*f(x) \tag{C.4}$$
$$> (Np_1)(1-F(Np_1)) + \sum_{0}^{Np_1} x*f(x) + (Np_2)(1-F(Np_2)) + \sum_{0}^{Np_2} x*f(x)$$

This ultimately simplifies to $p(n_2 \leq Np_2 - 1) > p(n_1 \leq Np_1)$ which are binomial CDFs. When the mean of a binomial distribution is an integer, then the median and mean coincide. Thus, under the simplifying assumption that $Np_1$ and $Np_2$ are integers[1]:

- $Np_2$ is a median $\Rightarrow p(n_2 \leq Np_2 - 1) < 0.5$
- $Np_1$ is a median $\Rightarrow p(n_1 \leq Np_1) \geq 0.5$
- $\Rightarrow 0.5 > p(n_2 \leq Np_2 - 1) > p(n_1 \leq Np_1) \geq 0.5$, which is a contradiction

Thus, it cannot be the case that an individual has an incentive to deviate from a report of $\{y_1, ..., y_C\} = E[\{n_1, ..., n_C\}|\{N; p_1, ..., p_C\}]$. □

Finally, note that we can transform participants' reports into probabilities. As shown in Claim 1, the mechanism incentives participants to report:

$$\{g_1, ..., g_C\} = E[\{n_1, ..., n_C\}|\{N; p_1, ..., p_C\}] = \{Np_1, ..., Np_C\} \tag{C.5}$$

Thus, $g_c/N = p_c$.

[1]This relationship holds without the integer assumption, but is seen more directly when mean and median coincide.

## Appendix D: Laboratory experiment instructions

*In the text below, instructions unique to Baseline treatment are contained in square brackets:* [...]. *Instructions unique to Visibility treatment are contained in curly brackets:* {...}. *All other instructions are identical across treatments.*

## Preliminary on-screen instructions

*(Note: All of the instructions in this section were displayed on a series of screens at participants' computer terminals and read aloud by the experimenter. The experimenter controlled the pace at which participants progressed through these screens.)*

***Introduction***

This experiment is a study of decision-making. You will receive $5 simply for showing up. You will have an opportunity for additional earnings depending on the decision that you make in two tasks: a giving task and a guessing task. In the giving task, you will be given $10 that you can contribute to charity. In the guessing task you can earn up to $7.

Please do not talk to other participants during the experiment. If at any point you have a question, raise your hand and we will come to you to answer it.

We will first introduce you to the charitable cause that you can donate to.

***Description of cause***

A dedicated team of local engineering professionals as well as University of Pittsburgh and Carnegie-Mellon students has been working on delivering potable water to the homes of a subsistence farming community living at 3600 meters elevation in the Andes Mountains of Ecuador.

However due to topography, lack of a continuous clean water supply has continue to plague development. The community-owned water source is a natural spring located 1000 feet below the village. Currently, community members rely on a daily 2 hour commute to collect potable water and supplement their needs through a rain catchment system constructed in 2009 by the Engineers Without Borders (EWB)-Pittsburgh team.

The Pittsburgh EWB team has lined up partial funding and community ownership for the infrastructure necessary for the ambitious project of providing a continuous water supply. Construction of the pipeline up the mountain is currently taking place. The last step of the project is to install faucets that directly deliver water to households. Your contribution will fund the cost of purchasing and installing faucets in Tingo Pucara.

As the world continues to grow larger and more complicated, there are people out there still walking hours every day to simply have clean water to drink, cook, and bathe. Help EWB-Pittsburgh cross Tingo Pucara off that list.

***Unconditional contribution instructions***

You will now have the opportunity to donate to the cause that was just described.

You have been given $10 in an envelope. Before proceeding, please sign the receipt that indicates that you have received $10.

You may contribute as much or as little of this money as you would like to the cause. However, we will ask that you restrict your contribution to $2 increments. That is, you will choose whether to contribute $0, $2, $4, $6, $8, or $10. Any money that you do not contribute is yours to keep.

[ For today's experiment, you have been randomly assigned to a group with two other participants in the room. Your group's total contributions will go towards purchasing and installing faucets in Tingo Pucara. You will not know the identity of your group members while deciding, nor will you be able to communicate with them about the decision.  Your group has the potential to make a large impact; the average cost of faucets and associated installation is $30, but varies with local conditions.

At the end of today's session, you will leave your donation in its original envelope on your desk.  The software will inform you of your group’s total donation to Tingo Pucara. ]

{ For today's experiment, you have been randomly assigned to a group with two other participants in the room. Your group's total contributions will go towards purchasing and installing faucets in Tingo Pucara. You will not know the identity of your group members while deciding, nor will you be able to communicate with them about the decision.  Your group has the potential to make a large impact; the average cost of faucets and associated installation is $30, but varies with local conditions.

At the end of today’s session, your group members will learn how much you contributed. In particular, you and your group (one group at a time) will go to a conference room with the experimenter to submit your contributions. While your group is waiting for its turn to go to the conference room, you and your group members will gather at the front of this room to fill out a group contribution slip together, indicating how much each member is donating. Your group members are the only participants who will observe how much you chose to give.

In the conference room, the experimenter will ask you for your group contribution slip. When the experimenter reads out your contribution from the slip, you wil count out your donation in front of your group members and hand it to the experimenter. At the end, the experimenter will inform you of your group’s total donation to Tingo Pucara and thank you for your donation.  }

In a moment the contribution screen will appear on your computer. Again, you may contribute any multiple of $2 in between $0 and $10. If you have a question, raise your hand and an experimenter will come to you to answer it.

*(Note: All remaining instructions were distributed as paper handouts and read aloud by an experimenter.)*

Instructions: Giving Task Part II

We will refer the decision you have just made as your Part I decision.

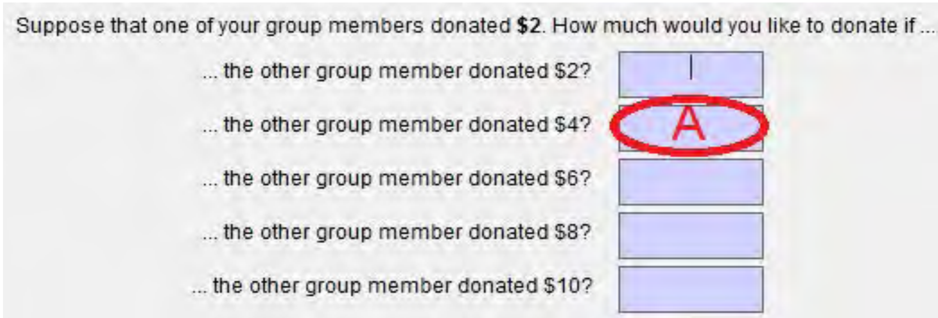
You may now have an opportunity to change your contribution based on what the other members of your group contributed.

Specifically, in a moment you will indicate what you would have contributed in Part I if you had observed your other two members' contributions. However, because you were not able to observe their contributions, you will be asked to indicate what your decision would have been given every possible pair of contributions.

To do so, you will be presented with a series of screens that capture every potential combination of your other group members' contributions from Part I. For example, the image below is the first screen that you will see; the input boxes in this screen represent all of the scenarios where one of your group members gave $0. Each line indicates the possible contributions of the *other* group member. In each input box, you will indicate the amount you would have contributed if your group members had contributed the amounts associated with that input box.


After you fill in the input boxes in the figure above, the screen below will show up. This screen represents all of the scenarios where one of your group members gave $2. (Notice that it leaves out the situation where one member gives $2 and the other gives $0. This is because you already entered a decision for that situation in the first screen.) For example, the input box labeled "A" is in the "$4" row. Therefore, in that input box you will enter the amount that you would want to contribute if you knew that one member contributed $2 and the other contributed $4.

After this you will see similar screens for all of the scenarios where one member gives $4, $6, $8, and finally $10.




**You will fill in *all* of the input boxes on each of these screens.** As before, you can enter any multiple of $2 from $0 to $10 in each input box.

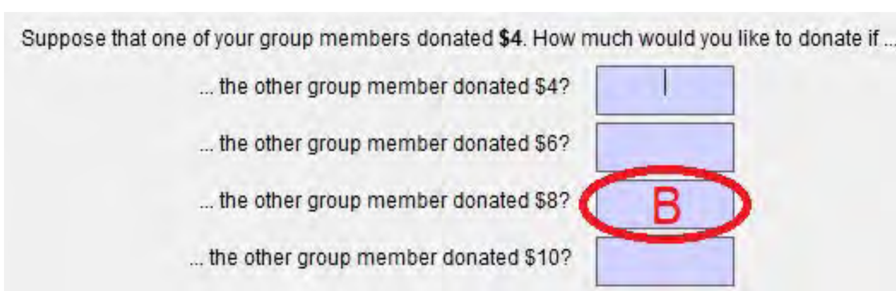
How might these decisions impact the contribution you actually provide?

At the end of the experiment, one member of each group will be randomly selected. If you are that person, the decisions that you make in this phase (Part II) will determine how much you actually contribute to the project.

**If you *are not* the randomly selected group member**, then the actual amount you will contribute at the end of the experiment will simply be the contribution you indicated in Part I of the giving task.

**If you *are* the randomly selected group member**, your contribution will be determined by your decision in this phase (Part II) conditional on your group members' contributions from Part I.

For example, suppose that at the end of the experiment you are informed that you are the randomly selected group member. This means that your two group members simply submit their contributions from Part I. Suppose that they chose $4 and $8. Because you are the randomly selected group member, your contribution will be whatever you specify in the input box associated with contributions of $4 and $8 (input box B in figure below.) Suppose you had entered $6 in that box. Then you would submit $6 and keep $4, regardless of the contribution you entered in Part I. Therefore, your group's total contribution would be $18 ($4+$8+$6).




You do not know whether you will be the randomly selected group member when you fill in the contribution tables. You will therefore have to think carefully about these decisions because they may determine how much you contribute to the project.

We are now ready to begin. Before proceeding to the contribution decisions, you will complete a brief quiz. This quiz has no impact on your earnings and is merely intended to ensure that the instructions are clear. Feel free to look back at the instructions while answering the questions. After each question you will be informed of the correct answer.

Instructions: Guessing task

You have now completed the giving task. We will now move to the guessing task, after which you will complete a brief survey. The software will then inform you of the following: 1) whether you were the randomly selected group member 2) your actual contribution to Tingo Pucara and 3) your earnings from the guessing task. This will conclude the session – you will then place your donation as indicated by the computer in its original envelope and leave it on your desk. [ The experimenter will then call your experimental ID to pay your show up fee and earnings from the guessing task. ] { This will conclude the session – the experimenter will then call each group to submit their donations. }

In the guessing task, you will guess the donation chosen **in Part I** by the other 14 participants in the room. Specifically, you have 14 tokens – one for each participant in the room (not including yourself). There were six possible contribution choices in Part I: $0, $2, $4, $6, $8, or $10. For each of these possible contributions, you will guess how many people in the room chose that contribution and assign your tokens accordingly. For example, if you think that everybody in the room chose $10 then you would assign all 14 tokens to "$10". If you think that half of the people in the room chose $6 and the other half chose $8, then you would assign 7 tokens to "$6" and 7 tokens to "$8."

For each token that you place correctly, you will receive $0.50. You receive nothing for each token that is placed incorrectly. For example, again suppose that you think that half of the participants chose $6 and the other half chose $8, so you assign 7 tokens to each of these. Suppose that instead, two people chose $6 and everyone else in the room chose $2. This means that exactly two of your guesses were correct so you would receive $1.00. If your guess had been correct – that is, if it were actually the case that 7 people chose $6 and 7 people chose $8 – then you would have received $7.

After everyone has completed the guessing task and a short survey, the software will display the information above. When you are informed of your earnings from the guessing task, please indicate this amount and show up fee on your receipt.