HARVARD | BUSINESS | SCHOOL

Incentives for Prosocial Behavior: The Role of Reputations

Christine Exley

Working Paper 16-063

Incentives for Prosocial Behavior: The Role of Reputations

Christine Exley
Harvard Business School

Working Paper 16-063

Copyright © 2015, 2016 by Christine Exley

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Incentives for Prosocial Behavior: The Role of Reputations

Christine Exley*

September 23, 2016

Abstract

Do monetary incentives encourage volunteering? Or, do they introduce concerns about appearing greedy and crowd out the motivation to volunteer? Since the importance of such image concerns is normally unobserved, the answer is theoretically unclear, and corresponding empirical evidence is mixed. To help counter this ambiguity, this paper proposes that the importance of image concerns – such as the desire to appear prosocial and not to appear greedy – relates to individuals' volunteer reputations. Experimental results support this possibility. Individuals with past histories of volunteering are less responsive to image concerns if their histories are public, or if their prosocial tendencies are already known. Consistent with a decreased importance of appearing prosocial, they are less likely to volunteer. Consistent with a decreased importance of not appearing greedy, they are less likely to be discouraged by public incentives.

JEL: C91; D64, H41

Keywords: incentives; image motivation; reputations; volunteer; prosocial behavior

*Exley: Harvard Business School, clexley@hbs.edu. Acknowledgements: I gratefully acknowledge funding for this study from the NSF (SES 1159032), Stanford's Institute for Research in the Social Sciences, and Stanford's Economics department.

1 Introduction

With over a quarter of Americans volunteering annually at an estimated market value of 175 billion dollars, understanding how to encourage volunteers to provide more help may yield significant benefits to crucial societal services (Warfield, 2013). A common strategy – offering small monetary incentives for volunteering – may backfire. For instance, monetary incentives may crowd out volunteers' intrinsic motivation.[1] Or, as suggested by Bénabou and Tirole (2006), public monetary incentives may crowd out volunteers' image motivation as they introduce concerns about appearing "greedy" instead of "prosocial." While experimental results in Ariely, Bracha and Meier (2009) support this possibility, it remains unclear in other settings when significant crowding out is likely to occur and thus limit the effectiveness of public incentives (see Table 1 for an overview of the mixed empirical literature).

To help overcome this ambiguity, this paper proposes that individuals' reputations, based on their past volunteer behavior, may play an important role. In particular, the importance of image concerns likely decreases when individuals' reputations are positive and public, or when others know more about their prosocial tendencies. In considering this possibility, with Bénabou and Tirole (2006) as a conceptual framework, three image effects about individuals' incentivized and public volunteer decisions follow.

First, absent public reputations, public incentives may introduce concerns about appearing greedy and thus result in significant crowd-out, as in Ariely, Bracha and Meier (2009). By varying the observability of the incentives to volunteer, this paper therefore tests for a *Negative Image Effect - public incentives, relative to private incentives, discourage public volunteer behavior.* By only manipulating the observability of the incentives (not the observability of the volunteer behavior or the level of the incentives), the Negative Image Effect excludes mechanisms other than the desire not to appear greedy that may cause crowding out to occur.

Second, even absent public incentives or concerns about appearing greedy, the observability of individuals' reputations may matter. If individuals' reputations about past volunteer behavior are public, choosing to volunteer may be less informative about their prosocial tendencies. By varying the observability of participants' reputations, this paper thus tests for a *Reputations Effect - public reputations, relative to private reputations, discourage public volunteer behavior.* In manipulating the observability of past volunteer behavior, while holding constant the observability of future volunteer behavior, two advantages arise. Unlike in observational data, where individuals are more likely to have public reputations if they volunteer more, this manipulation

[1] Crowding out of intrinsic motivation was first mentioned in Titmuss (1970), modeled in Bénabou and Tirole (2003), and argued in many empirical studies, such as Frey and Oberholzer-Gee (1997), Gneezy and Rustichini (2000), and Frey and Jergen (2001). Meier and Stutzer (2008) provide a nice discussion of various extrinsic and intrinsic motivations, and find that volunteers tend to be more intrinsically motivated, so crowd-out may be a particular concern. However, some later studies provide evidence against this crowding out, such as Goette and Stutzer (2008) and Ashaaf, Bandiera and Jack (2014).

facilitates the comparison of individuals with the same levels of past volunteer behavior. Additionally, while a robust finding in the literature is that more prosocial actions occur if they are more observable (Harbaugh, 1998*b*,*a*; Andreoni and Petrie, 2004; Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009; Ariely, Bracha and Meier, 2009; Lacetera and Macis, 2010*b*), the Reputations Effect considers how prosocial actions are influenced by prior – as opposed to current – observability conditions.[2] That is, the Reputations Effect considers whether there is a long-run downside to increased observability of volunteer behavior in so much as it crowds out future image motivation to volunteer.

Third, the interaction of public incentives with the observability of reputations may help predict whether individuals volunteer less because of concerns about appearing greedy. If individuals' reputations about past volunteer behavior are public, choosing to volunteer when provided with a public incentive may be less informative about the extent to which they are greedy. By examining how the observability of incentives interacts with the observability of reputations, this paper thus tests for an *Interactions Effect - public incentives, relative to private incentives, discourage public volunteer behavior less for those with public reputations than private reputations.* If the Interaction Effect holds, in addition to allowing for a better understanding as to when observability is likely to influence volunteer behavior, it may help to explain the mixed empirical findings on public incentives in the literature. Indeed, Table 1 shows that incentives appear more effective, and thus crowd-out concerns may have been less relevant, among study populations where reputations are likely more known.

The extent to which the three image effects influence volunteer decisions may vary according to an individual's type of reputation. For instance, individuals with better reputations, or individuals who have already exhibited prosocial tendencies, likely place a higher value on appearing prosocial.[3] The potential for a heterogenous effect based on past volunteer behavior aligns with growing evidence for consistency in prosocial tendencies. Gneezy et al. (2012) find that individuals who have engaged in costly prosocial behavior subsequently care more about appearing prosocial to themselves, Karlan and Wood (2014) observe more positive responses to information on aid effectiveness among larger previous donors, Exley (2016) shows that excuse-driven responses to risk are more likely among individuals who also give less when there is no risk in charitable giving decisions, and as detailed in Table 1, Niesse-Ruenzi, Weber and Becker (2014) and Lacetera, Macis and Slonim (2014) observe more positive responses to incentives among more frequent volunteers.

[2]Of course, more observability does not always lead to more prosocial behavior. For instance, this finding often does not result when other image concerns accompany individuals' actions, such as concerns about appearing "greedy" as in Ariely, Bracha and Meier (2009) or concerns related to the observability of one's income as in Bracha and Vesterlund (2013).

[3]A partially opposing possibility is that individuals with worse reputations, if they are public, may feel a greater need to overcome their poor reputations and thus volunteer more. This would only influence the Reputations Effect and Interactions Effect – a form of heterogeneity that is not consistent with the data in this study.

Results from a laboratory experiment with 130 participants provide qualitative support for the image effects. Results from an online Mturk study with 800 participants provide stronger support for the three image effects among individuals with histories of choosing to complete a previous volunteer activity (and thus "good" volunteer reputations) but not among individuals with histories of choosing not to complete a previous volunteer activity (and thus "bad" volunteer reputations). That is, for the group with likely stronger prosocial tendencies, public incentives discourage future volunteer behavior (the Negative Image Effect) and public reputations discourage future volunteer behavior (the Reputations Effect), but public reputations attenuate the extent to which public incentives discourage future volunteer behavior (the Interactions Effect).

Table 1: Literature Examining the Effect of Public Incentives on Public Volunteer Behavior

Paper	**Summary of relevant findings***
Mellström and Johannesson (2008)	In a field experiment with non-previous blood donors, incentives to complete health examinations to become blood donors discourage females and have no effect on males (- to 0 incentive response when reputations are likely unknown)
Gneezy and Rustichini (2000)	In a field experiment with school children, small monetary incentives to collect donations from the public have a negative effect and large monetary incentives have null effect (- to 0 incentive response when reputations are likely unknown to involved experimenter but known to children)
Ariely, Bracha and Meier (2009)	In a lab experiment with undergraduate students, incentivizing effort in a public volunteer task has no effect (0 incentive response when reputations are likely unknown to involved experimenter and most students)
Iajya et al. (2013)	In a field experiment with mostly non-previous blood donors, small supermarket vouchers have no effect and larger super market vouchers encourage more blood donations (0 to + incentive response when reputations are likely unknown among those at blood banks)
Carpenter and Myers (2010)	In observational data on firefighters, offering small stipends increases their turnout rate unless they have vanity license plates (0 to + incentive response when reputations are likely known among firefighters)
Lacetera, Macis and Slonim (2012)	In observational data and a field experiment with mostly previous blood donors, material incentives encourage more donations (+ incentive response when reputations are likely known among those at blood drives)
Niesse-Ruenzi, Weber and Becker (2014)	In a natural field experiment with new and previous blood donors, removing monetary compensation decreases blood donations, particularly for the most frequent blood donors (more + incentive response when reputations are likely more known)
Lacetera, Macis and Slonim (2014)	In a field experiment with previous donors, gift cards at a particular blood drive encourage more donations, even more so among those who have donated to that particular blood drive before, had donated more recently, and/or had donated more frequently (more + incentive response when reputations are likely more known)

*The classifications of the above study populations by reputations are highly speculative.

2 Online Study

In the online study, participants begin by completing a practice round of a simple task. This task requires participants to correctly count the number of 0s in seven different tables,

where each table contains a random series of fifty 0s and 1s.[4] Participants then make two decisions about whether to volunteer for the American Red Cross (ARC) by completing similar tasks. Instructions and required understanding questions precede each decision, and the study concludes with a short follow-up survey to gather demographic information and other controls for the analysis.

Before making their two decisions, image concerns are introduced by informing participants of their potential reward amounts. "Panel Members" (PMs), after observing some information on the participants' decisions, choose the reward amounts to be between \$0 and \$10. Participants know that any reward amount will be distributed to them as additional payments but will not influence the payments received by PMs.

For their first volunteer decisions, participants indicate whether they would like to volunteer for the ARC by completing the \$10-volunteer task. Completing the \$10-volunteer task requires a participant to solve seven tables and results in the ARC receiving a donation of \$10. Completing the \$10 volunteer task does not result in any additional payment for the participant. Prior to making this decision, participants know that PMs have a 50% chance of learning their \$10-volunteer task decisions and a 50% chance of not learning their \$10-volunteer task decisions.

For their second and now financially incentivized volunteer decisions, participants indicate whether they would like to volunteer for the ARC by completing the \$1-volunteer task. Completing the \$1-volunteer task requires a participant to solve seven tables and results in the ARC receiving a donation of \$1. Participants are also offered a financial incentive to volunteer; completing the \$1-volunteer task results in an additional \$1 for the participant by taking away \$1 from the PM. Given the anonymous online setting, the \$1 financial incentives come out of the PM's payment to bolster the saliency of and indeed provide a clear motive for caring about greedy image concerns. Prior to making this decision, participants learn exactly what information PMs will know before determining their reward amounts. This information varies on two dimensions.

First, participants learn the resolution of the 50% chance that their PMs will know their \$10-volunteer task decisions. Participants randomly selected to have public volunteer reputations (R_{pub}) learn that their PMs will know whether they chose to complete the \$10-volunteer task. By contrast, participants randomly selected to have private volunteer reputations (R_{priv}) learn that their PMs will not know whether they chose to complete the \$10-volunteer task.

Second, 50% chance (unbeknownst to participants) determines whether PMs know that participants are offered financial incentives to complete the \$1-volunteer task. Participants randomly assigned to the public incentive condition (I_{pub}) learn that their PMs will know the offered financial incentives to volunteer – i.e., will know how decisions influence the payments for the ARC, the participants, and the PMs. By contrast, participants randomly assigned to the private incentive condition (I_{priv}) learn that their PMs will not know that financial incentives to volunteer

[4]This is a modified version of the task employed by Abeler et al. (2011).

were offered – i.e., will only know how decisions influence the payments for the ARC.

In other words, when participants decide whether to complete the $1-volunteer task, they know their volunteer decisions will be public and are incentivized. Depending on their treatment condition, they also know whether their prior $10-volunteer task decisions will be public or private and whether their offered incentives to complete the $1-volunteer task will be public or private.[5] Table 2 summarizes this two-by-two design and Appendix A shows screenshots of the first decision about completing $10-volunteer task (same for all treatment groups) and the second decision about completing the $1-volunteer task (varies across the treatment groups).

Table 2: Treatment Groups

	Public Reputations (R_{pub})	**Private Reputations (R_{prvt})**
Public Incentives (I_{pub})	PM knows about incentives to volunteer and knows participant's volunteer reputation	PM knows about incentives to volunteer but does not know participant's volunteer reputation
Private Incentives (I_{prvt})	PM does not know about incentives to volunteer but knows participant's volunteer reputation	PM does not know about incentives to volunteer and does not know participant's volunteer reputation

A PM always knows a participant's decision to complete the incentivized $1-volunteer task. A participant's volunteer reputation indicates their decision to complete the $10-volunteer task.

As a final design note, elements of this design (and in particular the laboratory study detailed in Section 3) closely follow Ariely, Bracha and Meier (2009). The main difference involves the reputations variation and thus ability to test the Reputations Effect and Interactions Effect. The more subtle difference involves how the Negative Image Effect is tested. In Ariely, Bracha and Meier (2009), they show that incentives to volunteer (relative to no incentives) are less effective when everything is public than when everything is private. By instead separating the observability of incentives from the observability of volunteer behavior, this study tests for the Negative Image Effect by comparing the impact of public versus private incentives on volunteer behavior that is always public.[6]

[5]Importantly, one of the understanding questions for participants requires them to correctly indicate whether their PMs would learn: (i) what they chose in the first $10-volunteer task decision, (ii) what they chose in the second incentivized $1-volunteer task decision, and (iii) that they were offered a financial incentive of $1 (out of their PM's study compensation) to choose to complete the $1-volunteer task decision.

[6]This approach ensures the prosocial image and material incentives are constant across all treatment groups, and in doing so, prevents incentives from being ineffective due to a potential diminishing returns in terms of the "total incentive" offered. While no study, to my knowledge, has examined if there are diminishing returns such total incentives, Imas (2013) finds that individuals do not exert higher volunteer effort in response to higher charity payoffs, and Exley and Terry (2015) show that individuals may reduce their effort in response to higher charity payoffs due to reference-dependent behavior. The effect of incentives on charitable giving has also been extensively studied via estimating price elasticities of giving (Andreoni, 2006; Karlan and List, 2007).

2.1 Online Study Data and Implementation

This study was conducted via a Qualtrics survey on Amazon Mechanical Turk, a platform where "Mturk" workers can complete and receive payments for Human Intelligence Tasks (HITs). Eligible Mturk workers for this study include those who reside in the United States, have had at least 100 HITs approved, and have an approval rating of at least 95%. To learn more about experiments conducted on Amazon Mechanical Turk, including the replication of standard findings, see Paolacci, Chandler and Ipeirotis (2010) and Horton, Rand and Zeckhauser (2011).

A total of 800 Mturk workers completed this study as main participants and received the $4 completion fee on January 25th of 2016.[7] When completing this study, participants knew that there was a 1-in-50 chance that they would be a selected participant and have their decisions implemented. The selected participants had to complete the number of tables corresponding with their two decisions and corresponding payments were distributed as bonus payments. The non-selected participants did not have to complete any more tables and received no bonus payments.

An additional 20 Mturk workers completed a separate study as "Panel Members" (PMs) on January 26th and 27th of 2016.[8] They received a $3 completion fee, and if their randomly matched participant chose not to complete the $1-volunteer task, an additional $1 bonus payment.

Appendix Table A.1 shows that the resulting 183-220 participants in each treatment group did not differ on observable characteristics. 97% of participants were born in the United States, 56% percent have some college degree, and 53% are male. With an average of 15 volunteer hours in the past year, 72% report feeling favorably about the ARC and only 6% report feeling unfavorably about being offered incentives to volunteer.

2.2 Online Study Results

According to their willingness to complete the initial $10-volunteer task, 74% of participants enter the incentivized $1-volunteer task decision with "good" volunteer reputations and 26% enter with "bad" volunteer reputations. Figure 1 shows the corresponding impact of the treatment variations among these two groups.

To begin, notice that the minority (38%) of participants with bad volunteer reputations complete the incentivized $1-volunteer task, and these decisions are not significantly influenced by the treatment variations. This is consistent with participants who have bad volunteer reputations placing little value on image concerns.

By contrast, the majority (64%) of participants with good volunteer reputations complete the incentivized $1-volunteer task, and these decisions are influenced by the treatment variations. This is consistent with participants who have good volunteer reputations placing more value on appearing prosocial, and in particular, aligns with the three image effects as follows.

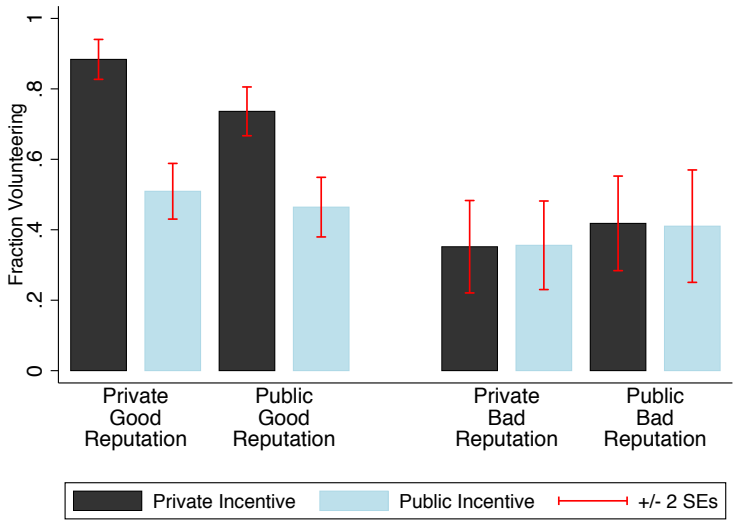
[7] This involved recruiting 801 Mturk workers due to 1 individual failing to complete the study.

[8] With 17 selected participants, 3 extra PMs reflect the randomization not yielding a perfect 1-1 match.

When both reputations and incentives to volunteer are private, 88% of these participants complete the $1-volunteer task. When reputations remain private but incentives to volunteer are instead public, the 88% volunteer rate significantly decreases to 51%. This finding supports the Negative Image Effect – public incentives discourage volunteering – and is consistent with concerns about appearing greedy crowding out volunteer behavior. When private incentives remain but volunteer reputations are instead public, the 88% volunteer rate significantly decreases to 74%. This finding supports the Reputations Effect – public reputations discourage volunteering – and is consistent with public reputations decreasing the importance of appearing prosocial.

In considering the interaction of the observability variations, notice that the extent of the Negative Image Effect, or drop in volunteer rates from public incentives, is about one-quarter smaller when reputations are public versus private (a drop of 28% versus 37%). This finding supports the Interactions Effect – public reputations attenuate the extent to which public incentives (relative to private incentives) discourage volunteering – and thus is consistent with public reputations limiting concerns about appearing greedy.

Figure 1: Fraction Choosing to Volunteer



This graph displays the fraction of individuals who choose to complete the incentivized $1-volunteer task across the four different treatment groups, separately for those with bad and good volunteer reputations.

To examine the statistical significance and robustness of the three image effects, I next consider results from Probit regressions of $\mathbb{P}(volunteer_i) = \Phi(\beta_0 + \beta_1 I_{pubi} + \beta_2 R_{pubi} + \beta_3 I_{pub} R_{pubi} + [Controls]_i)$, where $volunteer_i, I_{pubi}, R_{pubi}$, and $I_{pub} R_{pubi}$ are indicators for whether individual i volunteers, has a public incentive, has a public reputation, or has a public incentive and public reputation, respectively. While $\beta_1 < 0$ and $\beta_2 < 0$ provide direct support for the Negative Image Effect and Reputations Effect, respectively, $\beta_3 \neq 0$ does not necessarily imply support for the

Interactions Effect. $\beta_3 \neq 0$ shows that the interactions term helps to explain variation in the likelihood to volunteer, but the interpretation of an interactions term in a nonlinear model is more nuanced and discussed further below.

The first and second columns in Table 3 present results from Probit regressions when only controlling for whether a participant has a bad volunteer reputation (V_b=1) and when also controlling for a fuller set of demographic and belief controls, respectively. While there is significant support for the Negative Image Effect and Reputations Effect – both public incentives (I_{pub}) and public reputations (R_{pub}) discourage volunteering – the coefficients on the interaction terms of these variables are insignificant.

It may be important, however, to allow for differential treatment effects according to a participant's type of volunteer reputation. Indeed, as seen in Figure 1 and by the negative and significant coefficients on V_b, a participant with a bad volunteer reputation is less likely to volunteer – consistent with them placing less value on image concerns. The third column therefore interacts V_b with the treatment variables, and in addition to the coefficient on the interactions term (I_{pub}*R_{pub}) now becoming statistically significant, the support for the Negative Image Effect and Reputations Effect appears to strengthen. The fourth column shows that this set of results is further robust to controlling for measures of a participant's ability and reliability.

In considering the interpretation of these results, it is important to note that the distribution of marginal effects on interaction terms are not constant across covariates in nonlinear models (Ai and Norton, 2003). To thus narrow in on the interaction of interest – $I_{pub}R_{pubi}$ – while still allowing for heterogeneity by reputation, the fifth and sixth columns separately present results for participants with good reputations and bad reputations. Figure 2 then presents the corresponding distributions of estimated marginal effects (in the top row) and z-statistics (in the bottom row). While panel (b) shows that the estimates among individuals with bad reputations are never statistically significant, panel (a) shows that the the estimates are largely positive and marginally significant among individuals with good reputations. This supports the Interactions Effect holding among individuals with good reputations: the crowd out from public incentives is less likely if their past volunteer histories are public.

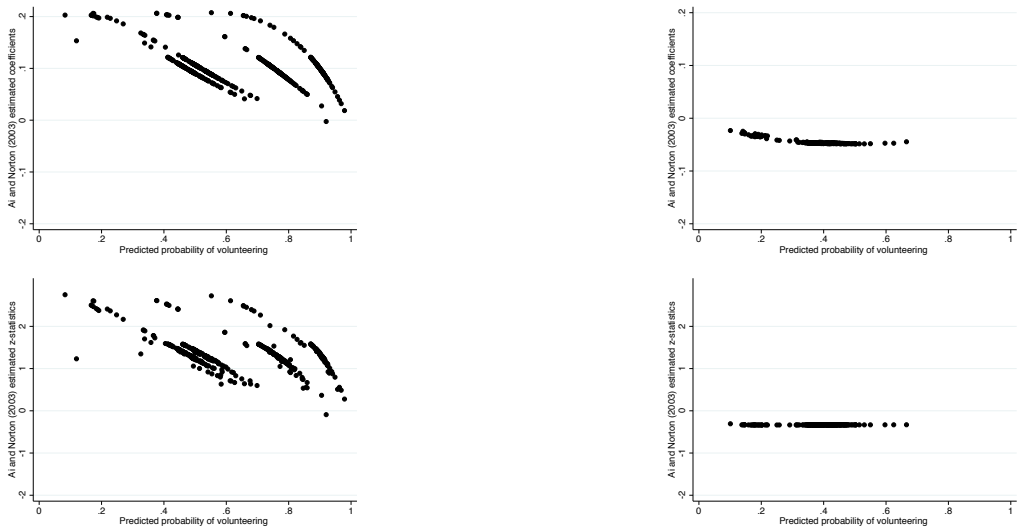
Table 3: Probit Regressions of Choice to Volunteer

	All				Reputations are:	
					Good	Bad
	(1)	(2)	(3)	(4)	(5)	(6)
I_{pub}	-0.76***	-0.78***	-1.22***	-1.23***	-1.23***	0.05
	(0.14)	(0.14)	(0.18)	(0.18)	(0.18)	(0.25)
R_{pub}	-0.26*	-0.27*	-0.58***	-0.59***	-0.60***	0.20
	(0.14)	(0.14)	(0.18)	(0.18)	(0.18)	(0.25)
I_{pub}*R_{pub}	0.19	0.18	0.47**	0.47**	0.48**	-0.12
	(0.19)	(0.19)	(0.23)	(0.23)	(0.24)	(0.37)
V^b	-0.73***	-0.76***	-1.63***	-1.64***		
	(0.11)	(0.11)	(0.23)	(0.23)		
I_{pub} *V^b			1.24***	1.25***		
			(0.30)	(0.30)		
R_{pub}*V^b			0.77**	0.79**		
			(0.31)	(0.31)		
I_{pub} *R_{pub}*V^b			-0.54	-0.54		
			(0.43)	(0.43)		
Constant	0.87***	0.52*	0.85***	0.57	0.50	-0.62
	(0.11)	(0.29)	(0.31)	(0.55)	(0.84)	(0.64)
Controls 1	no	yes	yes	yes	yes	yes
Controls 2	no	no	no	yes	yes	yes
Observations	800	800	800	800	593	207
Volunteer Rate	0.58	0.58	0.58	0.58	0.64	0.38

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results shown are from Probit Regressions of $\mathbb{P}(volunteer_i) = \Phi(\beta_0 + \beta_1 I_{pubi} + \beta_2 R_{pubi} + \beta_3 I_{pub} R_{pubi}[+\beta_4 V_i^b + \beta_5 I_{pub} V_i^b + \beta_6 R_{pub} V_i^b + \beta_7 I_{pub} R_{pub} V_i^b][+Controls_i])$. $volunteer_i$ is an indicator for whether individual i chooses to volunteer in the \$1-volunteer task. I_{pubi} and R_{pubi} are indicators for individual i having public incentives and public reputations, respectively. V_i^b is an indicator for individual i having a "bad" volunteer reputation from choosing not to volunteer in the \$10-volunteer task. Controls 1 include number of stated volunteer hours in past year and indicators for being male, being born in the US, having some college degree, feeling favorably about ARC, and feeling unfavorably about incentivized volunteering. In Column 6, not being born in the US perfectly predicts the dependent variable so that control is excluded. Controls 2 include practice round times and indicators for self-reports about the reliability of their decisions in the study.

Figure 2: Interaction Effect Results

(a) Good Reputation (b) Bad Reputation



Results are from estimates on the interaction term, $I_{pub}R_{pub}$, in the fifth and sixth columns of Table 3.

3 Laboratory Study

As in the online study, participants in the laboratory study decide how much to volunteer for American Red Cross (ARC) when their volunteer behavior is always public and incentivized. Depending on their treatment condition, participants also know whether their past volunteer behavior will be public or private and whether their offered incentives to volunteer will be public or private. The laboratory study differs from the online study in several ways, however.

First, instead of a binary decision about whether or not to volunteer, participants' incentivized volunteer decisions involve a more continuous measure of effort. In a task similar to that in Ariely, Bracha and Meier (2009), participants "click" or push a button on an electronic tally counter for eight minutes. For every five clicks a participant completes, the experimenter donates one cent to the ARC and adds one cent to a participant's study compensation.

Second, as participants' decisions are anonymous in the online study, image concerns are created by (i) forecasting that PMs will determine the participant's reward amount out of $10, and (ii) taking the $1 financial incentive to volunteer out of PMs' payments. While image concerns in the online study may then reflect a desire to be financially favored or not financially punished by others, the laboratory study offers a cleaner consideration of image concerns. In particular, the laboratory study does not forecast that PMs may determine participants' reward amounts (in this case, via a modified dictator game) and participants' incentives to volunteer do not influence their PMs' study compensation. Instead, image concerns arise from PMs being in the same laboratory study as their participants and thus being able to personally identify their

participants.[9]

Third, while the online study forms participants' reputations from their first decision about whether to complete the $1-volunteer task, the laboratory study forms participants' reputations largely from provided information on past volunteer behavior. In particular, participants report how many hours they have volunteered in the past year, and decide whether or not to complete a short volunteer survey task at the end of this study.[10] Each participant is then labeled as an "above average volunteer" if they volunteered 23 or more hours in the past year (the national average among young adults) or as a "below average volunteer" otherwise.[11] A key advantage of this constructed reputation is that it may be more informative about a participant's overall prosocial tendencies, given that it focuses on volunteer behavior in the past year as opposed to in one task. A potential downside to this constructed reputation is that participants could have misreported their past volunteer behavior. For the purposes of this study, however, a more related concern involves whether participants believe the PMs find their reported volunteer behavior to be informative about the extent to which they are prosocial. In support of this less demanding condition, Gneezy (2005) shows that participants who may financially benefit from lying to others tend to believe others will believe they are telling the truth. An additional reputations verification study indeed supports that participants believe PMs will find their reputation informative.

Fourth, while participants have a 1-in-50 chance of their volunteer decisions being implemented in the online study, all participants' volunteer decisions are implemented in the laboratory study.

Fifth, while participants are randomly assigned to one of the four treatment groups in both studies, the assignment procedure differs slightly in the laboratory study. The observability of the incentive offer differs on the session-level while the observability of reputations is randomly determined on the participant-level.

3.1 Laboratory Study Data and Implementation

This study was conducted at the Stanford Economics Research Laboratory. A total of 168 undergraduate students from Stanford University participated in the study between January and March 2012. In particular, the 130 main participants (4 participants are excluded for clicking incorrectly and 34 participants were PMs) led to the following results.[12] Appendix Table A.2 shows that the randomly assigned groups of participants did not differ on observable characteristics. 78% of participants were born in the United States, 12% percent are economics majors,

[9] Each participant's decision is observed by two PMs.

[10] Participants were informed that this short volunteer survey task would take approximately 5-10 minutes, would be given on behalf of Stanford's Haas Center for Public Service, and could be completed immediately after their participation in this study or later via a link sent to them by email.

[11] This cutoff of 23 hours was determined by calculating the average volunteer hours among young adults (16-24) from the Corporation for National & Community Service 2010 data about volunteering in America.

[12] Participants were instructed to hold the electronic tally counter in only one self-chosen hand and to only use their thumb to push the button. Excluded participants did not do this.

and 52% are male. With an average of 71 volunteer hours in the past year, 77% report feeling favorably about the ARC and only 24% report feeling unfavorably about being offered incentives to volunteer.

3.2 Laboratory Study Results

Similar to the online study, I consider regression results from $volunteer_i = \beta_0 + \beta_1 I_{pubi} + \beta_2 R_{pubi} + \beta_3 I_{pub} R_{pubi} + [Controls]_i$, where $volunteer_i$ equals how much individual i volunteers or "clicks", and I_{pubi}, R_{pubi}, and $I_{pub} R_{pubi}$ are indicators for whether individual i has a public incentive, has a public reputation, or has a public incentive and public reputation, respectively. Given this, $\beta_1 < 0$ supports the Negative Image Effect, $\beta_2 < 0$ supports the Reputations Effect, and $\beta_3 \neq 0$ supports the Interactions Effect. Importantly, recall that any support for these image effects arises from participants' concerns about how they appear to the PMs absent a financial reason to care about how PMs view them. That is, the laboratory results serve as a useful complement to the anonymous online study, where image concerns may instead reflect a desire to be financially favored by the PMs.

The first and second columns in Table 4 present qualitatively consistent, but not statistically significant, results from Ordinary Least Squares regressions when only controlling for whether a participant has a "bad" volunteer reputation from having below average volunteer hours in the past year (V_b=1) and when also controlling for a fuller set of demographic and belief controls, respectively. The third column presents similar results when further allowing for differential treatment effects by reputation type; the interactions of V_b with the treatment variables are also not statistically significant. That is, unlike with the online study, evidence for the image effects and evidence for heterogeneous effects by the type of volunteer reputation is underpowered.[13]

While there is not significant support for the image effects at the average level of volunteering, a closer look at the distributions of volunteer behavior is suggestive. As shown by the quantile regressions in the fourth, fifth and sixth columns of Table 4, there is statistically significant support for the image effects at the median and 75th percentile. This is consistent with the image effects mostly influencing those who may be relatively more image concerned – i.e., the top portion of the distribution or the "top volunteers."

[13] Although not shown, there is qualitatively suggestive evidence for females responding more strongly to the image effects. This relates to prior literature; for instance, Mellström and Johannesson (2008) and Lacetera and Macis (2010*a*) find that females are particularly averse to monetary incentives for completing a health examination to become a blood donor and for donating blood, respectively. More broadly, Croson and Gneezy (2009) provide a survey of the literature and suggest that women may be "more sensitive to subtle cues than" men, which is further supported by findings such as in Andreoni and Vesterlund (2001), DellaVigna et al. (2013) and Jones and Linardi (2014).

Table 4: Regressions of Volunteer Effort - i.e., Clicks

	OLS Regressions			Quantile Regresions		
				25%-ile	50%-ile	75%-ile
	1	2	3	4	5	6
I_{pub}	-65.94	-71.25	-62.74	-21.50	-159.64**	-210.50***
	(59.48)	(59.45)	(74.94)	(107.39)	(68.04)	(75.08)
R_{pub}	-83.99	-81.06	-75.66	43.50	-188.51***	-185.69***
	(55.20)	(55.58)	(67.62)	(100.39)	(63.61)	(70.18)
I_{pub}*R_{pub}	46.28	52.71	76.32	-43.56	212.75**	207.99**
	(81.59)	(82.43)	(97.98)	(148.90)	(94.34)	(104.10)
V^b	21.97	-24.74	10.05	-81.75	-76.46	49.69
	(44.78)	(53.91)	(86.54)	(97.38)	(61.70)	(68.08)
I_{pub}*V^b			-15.65			
			(128.54)			
R_{pub}*V^b			-6.38			
			(120.26)			
I_{pub}*R_{pub}*V^b			-134.16			
			(192.42)			
Constant	2259.34***	2188.46***	2185.34***	2078.72***	2307.39***	2411.42***
	(43.93)	(93.67)	(96.39)	(169.20)	(107.20)	(118.29)
Controls	no	yes	yes	yes	yes	yes
N	130	130	130	130	130	130
X%-ile/avg clicks	2202.46	2202.46	2202.46	2054.00	2209.00	2349.00

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The first three columns involve OLS regressions. The last three columns involves quantile regressions on the 25th, 50th, and 75th percentile. All regression results are of $volunteer_i = \beta_0 + \beta_1 I_{pubi} + \beta_2 R_{pubi} + \beta_3 I_{pub} R_{pubi} (+\beta_4 V_i^b + \beta_5 I_{pub} V_i^b + \beta_6 R_{pub} V_i^b + \beta_7 I_{pub} R_{pub} V_i^b)(+Controls_i) + \epsilon_i$. $volunteer_i$ equals how much individual i volunteers or "clicks." I_{pubi} and R_{pubi} are indicators for individual i having public incentives and public reputations, respectively. V_i^b is an indicator for individual i having a "bad" volunteer reputation from having a below average number of volunteer hours in the prior year. Controls include number of stated volunteer hours in past year and indicators for deciding to complete the volunteer survey task, being male, being born in the US, being an economics major, feeling favorably about ARC. and, feeling unfavorably about incentivized volunteering.

4 Conclusions and Discussion

This paper tests three image effects: whether public incentives discourage volunteering (the Negative Image Effect), whether public reputations discourage volunteering (the Reputations Effect), and whether public reputations attenuate the extent to which public incentives discourage volunteering (the Interactions Effect). While the laboratory study provides qualitative support for the three image effects, the online study provides stronger support among individuals with prior volunteer behavior or "good" volunteer reputations but not among individuals without prior volunteer behavior or "bad" volunteer reputations. The Interactions Effect in particular may help to unify the mixed literature on incentivizing volunteer behavior: public incentives to volunteer are less likely to be hampered by a desire to avoid appearing "greedy" among individuals with good and public, rather than private, volunteer reputations.

Future work may investigate whether non-profit organizations can capitalize on these findings. All else equal, private incentives may be preferable to public incentives since they do not introduce concerns about appearing greedy. Private incentives may often be costly to implement though, as they cannot be shared at large events or prominently displayed online. When targeting individuals with private volunteer reputations rather than public volunteer reputations, however, the Interactions Effect considered in this paper suggests that nonprofit organizations should be more willing to incur the corresponding implementation costs of private incentives.

Future work may also delve into whether it is beneficial for nonprofit organizations to promote public volunteer reputations – via award ceremonies or other public recognition tactics. While a robust finding in the prosocial behavior literature involves more prosocial actions arising when they are more observable, the Reputations Effect documented in this paper highlights the downside to past actions being observable. More observability may increase prosocial actions today but decrease prosocial actions tomorrow. When considering this potential tradeoff, nonprofit organizations may therefore reach different conclusions depending on their desired outcomes, such as whether they need to engage individuals in a single volunteer activity or repeated volunteer activities.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101: 470–492.

Ai, Chunrong, and Edward C. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters*, 80(1): 123–129.

Andreoni, James. 2006. "Chapter 18 Philanthropy." In *Applications*. Vol. 2 of *Handbook on the Economics of Giving, Reciprocity and Altruism*, , ed. Serge-Christophe Kolm and Jean Mercier Ythier, 1201 – 1269. Elsevier.

Andreoni, James, and B. Douglas Bernheim. 2009. "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*, 77(5): 1607–1636.

Andreoni, James, and Lise Vesterlund. 2001. "Which is the fair sex? Gender differences in altruism." *Quarterly Journal of Economics*, 116(1): 293–312.

Andreoni, James, and Ragan Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fund-raising." *Journal of Public Economics*, 88: 1605–1623.

Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review*, 99(1): 544–555.

Ashaaf, Nava, Oriana Bandiera, and Kelsey Jack. 2014. "No margin, no mission? A Field Experiment on Incentives for Pro-Social Tasks." *Journal of Public Economics*, 120: 1–17.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies*, 70(3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–1678.

Bracha, Anat, and Lise Vesterlund. 2013. "How Low Can You Go? Charity Reporting When Donations Signal Income and Generosity." *FRB Boston Working Papers Series, paper no. 13-11.*

Carpenter, Jeffrey, and Caitlin Knowles Myers. 2010. "Why volunteer? Evidence on the role of altruism, image, and incentives." *Journal of Public Economics*, 94(11-12): 911 – 920.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2): 448–474.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2013. "The Importance of Being Marginal: Gender Differences in Generosity." NBER Working Paper No. 18748.

Exley, Christine L. 2016. "Excusing Selfishness in Charitable Giving: The Role of Risk." *Review of Economic Studies*, 83(2): 587–628.

Exley, Christine L., and Stephen J. Terry. 2015. "Wage Elasticities in Working and Volunteering: The Role of Reference Points in a Laboratory Study." *Working Paper*.

Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out." *The American Economic Review*, 87(4): pp. 746–755.

Frey, Bruno S., and Reto Jergen. 2001. "Motivaiton of Crowding Theory." *Journal of Economic Surveys*, 15(5): 589–611.

Gneezy, Ayelet, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton. 2012. "Paying to Be Nice: Consistency and Costly Prosocial Behavior." *Management Science*, 58(1): 179–187.

Gneezy, Uri. 2005. "Deception: The Role of Consequences." *The American Economic Review*, 95(1): pp. 384–394.

Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics*, 115(3): 791–810.

Goette, Lorenz F., and Alois Stutzer. 2008. "Blood Donations and Incentives: Evidence from a Field Experiment." *SSRN eLibrary*.

Harbaugh, William T. 1998*a*. "The Prestige Motive for Making Charitable Transfers." *The American Economic Review*, 88(2): pp. 277–282.

Harbaugh, William T. 1998*b*. "What do donations buy?: A model of philanthropy based on prestige and warm glow." *Journal of Public Economics*, 67(2): 269 – 284.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The online laboratory: Conducting experiments in a real labor marke." *Experimental Economics*, 14(3): 399–425.

Iajya, Victor, Nicola Lacetera, Mario Macis, and Robert Slonim. 2013. "The Effects of Information, Social and Economic Incentives on Voluntary Undirected Blood Donations: Evidence from a Randomized Controlled Trial in Argentina." *Social Science and Medicine*, 98: 214–233.

Imas, Alex. 2013. "Working for the "warm glow": On the benefits and limits of prosocial incentives." *Journal of Public Economics*.

Jones, Daniel, and Sera Linardi. 2014. "Wallflowers Doing Good: Field and Lab Evidence of Heterogeneity in Reputation Concerns." *Management Science*.

Karlan, Dean, and Daniel H. Wood. 2014. "The effect of effectiveness: donor response to aid effectiveness in a direct mail fundraising experiment." *NBER Working Paper Series, Working Paper 20047*.

Karlan, Dean, and John A. List. 2007. "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment." *The American Economic Review*, 97(5): pp. 1774–1793.

Lacetera, Nicola, and Mario Macis. 2010*a*. "Do all material incentives for pro-social activities backfire? The response to cash and non-cash incentives for blood donations." *Journal of Economic Psychology*, 31(4): 738 – 748.

Lacetera, Nicola, and Mario Macis. 2010*b*. "Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme." *Journal of Economic Behavior and Organization*, 76(2): 225 – 237.

Lacetera, Nicola, Mario Macis, and Robert Slonim. 2012. "Will There Be Blood? Incentives and Displacement Effects in Pro-Social Behavior." *American Economic Journal: Economic Policy*, 4(1): 186–223.

Lacetera, Nicola, Mario Macis, and Robert Slonim. 2014. "Rewarding Volunteers: A Field Experiment." *Management Science*, 1–23.

Meier, Stephan, and Alois Stutzer. 2008. "Is Volunteering Rewarding in Itself?" *Economica*, 75: 39–59.

Mellström, Carl, and Magnus Johannesson. 2008. "Crowding out in Blood Donation: Was Titmuss Right?" *Journal of the European Economic Association*, 6(4): 845–863.

Niesse-Ruenzi, Alexandra, Martin Weber, and David Michael Becker. 2014. "To pay or not to pay – Evidence from whole blood donations in Germany." *Working paper*.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. "Running experiments on amazon mechanical turk." *Judgment and Decision Making*, 5(5): 411–419.

Titmuss, Richard M. 1970. *The Gift Relationship.* London:Allen and Unwin.

Warfield, Samantha Jo. 2013. "New Federal Report Finds 1 in 4 Americans Volunteer." *Corporation for National & Community Service [Press Release].*

A Appendix

Figure A.1: Screenshot of first decision about $10-volunteer task

Decision 1: If you are selected, would you like to volunteer for the American Red Cross (ARC) by completing the $10-volunteer task?

◯ **Yes - I would like to complete the $10-volunteer task**

Before deciding your reward amount (out of $10), your panel member would then have a 50% chance of learning that you VOLUNTEERED and thus earned $10 for the ARC.

◯ **No - I would NOT like to complete the $10-volunteer task**

Before deciding your reward amount (out of $10), your panel member would then have a 50% chance of learning that you DID NOT VOLUNTEER and thus earned $0 for the ARC.

Figure A.2: ($I_{priv}R_{priv}$, regardless of first decision) Screenshot of second decision about incentivized $1-volunteer task

Decision 2: If you are selected, would you like to volunteer for the American Red Cross (ARC) by completing the $1-volunteer task?

◯ **Yes - I would like to complete the $1-volunteer task (and thus accept the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: UNKNOWN and thus earned unknown amount for ARC
- Decision 2: VOLUNTEERED and thus earned $1 for ARC

◯ **No - I would NOT like to complete the $1-volunteer task (and thus decline the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: UNKNOWN and thus earned unknown amount for ARC
- Decision 2: DID NOT VOLUNTEER and thus earned $0 for ARC

Figure A.3: ($I_{priv}R_{pub}$, if chose to volunteer in first decision) Screenshot of second decision about incentivized $1-volunteer task

Decision 2: If you are selected, would you like to volunteer for the American Red Cross (ARC) by completing the $1-volunteer task?

○ **Yes - I would like to complete the $1-volunteer task (and thus accept the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1:VOLUNEERED and thus earned $10 for ARC
- Decision 2: VOLUNTEERED and thus earned $1 for ARC

○ **No - I would NOT like to complete the $1-volunteer task (and thus decline the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: VOLUNTEERED and thus earned $10 for ARC
- Decision 2: DID NOT VOLUNTEER and thus earned $0 for ARC

Figure A.4: ($I_{pub}R_{priv}$, regardless of first decision) Screenshot of second decision about incentivized $1-volunteer task

Decision 2: If you are selected, would you like to volunteer for the American Red Cross (ARC) by completing the $1-volunteer task?

○ **Yes - I would like to complete the $1-volunteer task (and thus accept the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

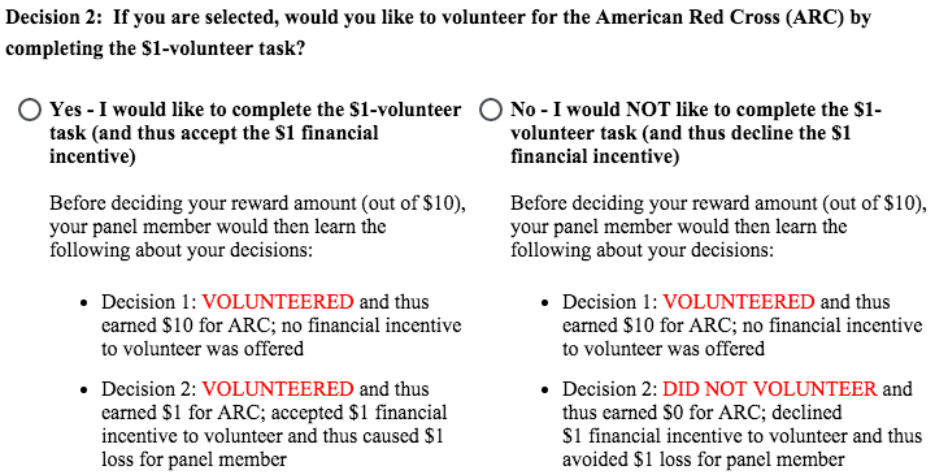
- Decision 1: UNKNOWN and thus earned unknown amount for ARC; no financial incentive to volunteer was offered
- Decision 2: VOLUNTEERED and thus earned $1 for ARC; accepted $1 financial incentive to volunteer and thus caused $1 loss for panel member

○ **No - I would NOT like to complete the $1-volunteer task (and thus decline the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: UNKNOWN and thus earned unknown amount for ARC; no financial incentive to volunteer was offered
- Decision 2: DID NOT VOLUNTEER and thus earned $0 for ARC; declined $1 financial incentive to volunteer and thus avoided $1 loss for panel member

Figure A.5: ($I_{pub}R_{pub}$, if chose to volunteer in first decision) Screenshot of second decision about incentivized $1-volunteer task



Decision 2: If you are selected, would you like to volunteer for the American Red Cross (ARC) by completing the $1-volunteer task?

○ **Yes - I would like to complete the $1-volunteer task (and thus accept the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: VOLUNTEERED and thus earned $10 for ARC; no financial incentive to volunteer was offered
- Decision 2: VOLUNTEERED and thus earned $1 for ARC; accepted $1 financial incentive to volunteer and thus caused $1 loss for panel member

○ **No - I would NOT like to complete the $1-volunteer task (and thus decline the $1 financial incentive)**

Before deciding your reward amount (out of $10), your panel member would then learn the following about your decisions:

- Decision 1: VOLUNTEERED and thus earned $10 for ARC; no financial incentive to volunteer was offered
- Decision 2: DID NOT VOLUNTEER and thus earned $0 for ARC; declined $1 financial incentive to volunteer and thus avoided $1 loss for panel member

Table A.1: By Observability of Incentives and Reputations: Characteristics of Participants in Online Study

	All	R_{priv}, I_{priv}	R_{priv}, I_{pub}	R_{pub}, I_{priv}	R_{pub},I_{pub}
Bad Volunteer History (V^b)	0.26	0.30	0.27	0.25	0.22
Born in United States	0.97	0.96	0.96	0.97	0.98
Male	0.53	0.50	0.48	0.56	0.56
Some College Degree	0.56	0.60	0.55	0.57	0.53
Favorable about ARC	0.72	0.70	0.68	0.78	0.70
Unfavorable about Incentives	0.06	0.07	0.05	0.06	0.06
Volunteer Hours	15.88	18.14	14.50	15.87	15.30
N	800	183	220	218	179

All of the above values indicate the fraction of participants with a given characteristic, except for the values associated with Volunteer Hours which indicate the average stated past volunteer hours. Participants were randomly assigned to one of the four treatment groups. Out of the six pairwise comparisons between any two treatment groups, I can never reject any joint hypothesis that the means of the variables are the same. Out of the forty-two pairwise comparisons for a given variable across any two treatment groups, I can only reject one individual hypothesis that the means of a variable are the same at the 0.10 significance level.

Table A.2: By Observability of Incentives and Reputations: Characteristics of Participants in Laboratory Study

	Full Subsample	R_{pub}	R_{prvt}	I_{pub}	I_{prvt}
Below Average Volunteer (V^b)	0.30	0.25	0.36	0.25	0.34
Born in United States	0.78	0.80	0.77	0.85	0.73
Male	0.52	0.49	0.54	0.53	0.51
Economics Major	0.12	0.12	0.11	0.10	0.13
Favorable about ARC	0.77	0.75	0.79	0.81	0.73
Unfavorable about Incentives	0.24	0.23	0.25	0.27	0.21
Volunteer Hours (in Past Year)	70.90	75.32	65.90	81.49	62.10
Completed Volunteer Survey Task	0.74	0.70	0.79	0.76	0.72
N	130	69	61	59	71

All of the above values indicate the fraction of participants with a given characteristic, except for the values associated with Volunteer Hours which indicate the average volunteer hours for participants. Across (1) the R_{pub} and R_{prvt} subsamples, and (2) the I_{pub} and I_{prvt} subsamples, I cannot reject that joint hypothesis that the means of the above variables are different, and I also cannot reject any individual hypothesis that the mean of any of the above variables are different.