Spring 2021

Three Essays on the Economics of Philanthropy

Yuxin Su
Claremont Graduate University

Three Essays on the Economics of Philanthropy

By

Yuxin Su

Claremont Graduate University

2021

Approval of the Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Yuxin Su as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Economics.

C. Mónica Capra, Chair
Claremont Graduate University
Professor of Economic Sciences

Joshua Tasoff
Claremont Graduate University
Associate Professor of Economic Sciences

Yi Feng
Claremont Graduate University
Luther Lee Jr. Memorial Chair in Government

Abstract

Three Essays on the Economics of Philanthropy

By

Yuxin Su

Claremont Graduate University: 2021

This dissertation studies the economics of prosocial behavior. More specifically, I investigate experimentally how prosocial incentives, pledges, and altruistic self-concept affect individuals' prosocial behaviors in three chapters.

Chapter 1 studies the role of self-chosen goals in shaping prosocial incentives' motivation on individuals' performance. The prosocial incentive is a way of motivation where workers' payments are associated with an additional reward to the charities. It is now widely accepted by the firms because it helps build the corporate culture, and boosts employee's morale, performance, and job satisfaction. However, recent studies have shown that a larger reward size does not necessarily increase workers' prosocial incentives. To solve this limitation, I implement a self-chosen goal scheme along with the incentives. I design an online experiment in which participants set goals for themselves engaging in real effort tasks. Participants obtain prosocial rewards only when they reach their goals. My results show that workers who receive prosocial incentives improve their performance by setting higher goals and achieving them. Moreover, when provided with the opportunity to receive large rewards, workers who are matched with the charity's mission will set higher goals to motivate themselves further to make additional efforts. My findings suggest prosocial incentives are comparable to monetary incentives in motivating workers within a self-chosen goal scheme. The preferred type of incentive depends on the firm's target and worker's heterogeneity.

Chapter 2 investigates experimentally whether pledges with respect to when one volunteer increase volunteering. As shown in previous literature, the effect of pledging on volunteering is ambiguous. On one hand, pledges can boost volunteering as it offers volunteers the option to choose when to help others. On the other hand, pledges open the doors for individuals to find more ways to excuse themselves from having to volunteer. In this paper, we study how volunteering decisions are affected by pledges using an online experiment. We find that pledges increase reneging on promises to volunteer, but total effort donations do not change. We also develop a simple model that helps explain the ways in which relevant parameters affect behaviors in our experiment. In particular, our model predicts that when given the opportunity to pledge to volunteer, people with high altruism or high warm-glow prefer to volunteer sooner rather than later, while higher future expected participation costs and lower expected reneging costs result in lower rates of rejection immediately. Moreover, pledges increase reneging behaviors on the future date, because those who want to volunteer don't delay their volunteering; however, those whose costs of saying "no" are high, are driven to postpone their rejection and renege on the future date.

Chapter 3 digs deeper to study the effect of personality traits on the willingness to make and keep a promise to volunteer. In our experiment, Amazon Mechanical Turk participants are given the option to volunteer by donating time and effort to a charity. They also answer a series of questionnaires, including the Big Five personality test and attitudinal questions that we use to construct an index representing altruistic self-concept. Self-concept refers to the way we describe and evaluate ourselves. We find that altruistic self-concept mediates how personality affects volunteering decisions. In particular, agreeableness has a strong influence on the probability of making and keeping promises to volunteer through its effect on altruistic self-concept. Our findings have useful implications for non-profit organizations. Agreeable individuals who evaluate and describe themselves as altruistic can be more helpful and dependable, so that the organizations can find ways to strengthen altruistic self-concept, thereby positively influencing prosociality in the workplace.

Dedication

This dissertation is dedicated ...

To my parents, family

and to my feline friend

Acknowledgements

I would like to express my sincere gratitude and pay special appreciation to the following people, who generously contribute to the research presented in this dissertation and assisted me during my doctoral studies.

First and foremost, I am incredibly grateful to my advisor Professor C. Mónica Capra for her invaluable guidance, vital support, immense knowledge, and continuous encouragement during my years at Claremont Graduate University. She has always been an incredible teacher and advisor.

I am highly indebted to the other members of my dissertation committee Professor Joshua Tasoff and Professor Yi Feng, for their insightful suggestions and advice, and for providing valuable comments and advice on my research. I would also like to express my sincere gratitude to Professor Nancy Neiman for her guidance on effective teaching.

This dissertation benefits greatly from discussion with colleagues and seminar participants at CGU and the external conferences participants. I sincerely appreciate the support from my wife Yu Cao, as well as my colleagues and friends Ning Li, Kichang Kim, Bobae Hong, Guan Wang, Chengcheng Zhang, Jin Xu, Bing Jiang, Yanyan Yang, Jae Joon Lee, and Hae Won Byun.

Last but not least, I would like to thank the current and former administrative staff of the Economics Department. Lidia Romero, Priscilla Elizarraras, Eliana Leon, and Lauren Copeland have always been patient and provide various administrative support throughout my years at CGU.

Contents

1 Chapter 1: Can Prosocial Incentives within a Goal-Setting Scheme Motivate Workers?

A larger reward size does not necessarily increase workers' prosocial incentives. This paper shows that under a self-chosen goal scheme, there does exist a strong positive association between these two. We design an online experiment in which participants set goals for themselves engaging in real effort tasks. Participants obtain prosocial rewards only when they reach their goals. Our results show that workers who receive prosocial incentives improve their performance by setting higher goals and achieving them. Moreover, when provided with the opportunity to receive large rewards, workers who are matched with the charity's mission will set higher goals to motivate themselves further to make additional efforts. Our findings suggest prosocial incentives are comparable to monetary incentives in motivating workers within a self-chosen goal scheme. The preferred type of incentive depends on the firm's target and worker's heterogeneity.

1.1 Introduction

Prosocial rewards, additional compensation made by firms in forms of charitable contribution, would increase workers' performance. Providing prosocial incentives in workers' compensation is now a prevalent method among companies to motivate their workers. According to the data collected by the non-profit organization Accounting Principals (2017), 38% of companies provided their employees with prosocial rewards. For example, Credit Suisse requires its employees to give 2.5% of their bonuses to charity. In recent studies, economists, psychologists, and sociologists have studied why prosocial incentive motivation is widely accepted.

There are two ways in which prosocial incentives benefit workers. First, prosocial rewards enable workers to establish altruistic self-image about themselves and obtain self-fulfillment. Second, prosocial rewards motivate workers to provide more effort, which eventually increases their compensation (e.g., when workers are paid by piece rate). Therefore, workers are much

happier and have higher satisfaction in jobs with prosocial rewards than those with monetary rewards and those without any rewards (Anik et al., 2013; Berger, Guo, and Presslee, 2018). In a recent study, O'Brien and Kassirer (2019) found that happiness decreases dramatically when people are consistently exposed to similar physical rewards. However, happiness would last longer when people are helping others. Their findings indicate that prosocial incentives could increase workers' job satisfaction and such impact is long-lasting.

Prosocial incentives also benefit firms. First of all, firms can establish prosocial reputations from their charitable donation and develop a culture of "meaning, purpose and belonging" (Lottinville, 2018). This is one of the reasons that corporate social responsibility (CSR) has become an indispensable aspect of the business. Moreover, firms' charitable donations bring them tax benefits (Publication 535, 2018). Furthermore, prosocial rewards increase firms' profit by improving workers' performance (Tonin and Vlassopoulos, 2015).

Despite its current prevalence, recent studies have pointed out some limitations of applying prosocial incentives in increasing workers' performance. Different from monetary rewards, a larger prosocial reward size does not necessarily increase a worker's performance. That is, there is a non-monotonic relationship between prosocial reward size and worker's performance (Charness, Cobo-Reyes, and Sanchez, 2016; K. K. Li, 2019; Imas, 2014; Khan, Goldsmith,and Dhar, 2020; DellaVigna and Pope, 2018). A plausible explanation for this phenomenon is "warm-glow giving" effect (Andreoni, 1989; Andreoni, 1990). That is, when an individual works for prosocial purposes, she only cares about whether she helps others or not, rather than how much help she gives. Thus, workers' output level will not monotonically increase with the prosocial reward size.

Workers' reduced professionalism under prosocial incentive is another limitation (Cassar and Meier, 2017; List and Momeni, 2017; K. K. Li, 2019). For example, K. K. Li (2019) designed an accuracy based payment in a lab experiment and showed that comparing to workers did not know the nature of their tasks, workers that knew their task were for charities would work less carefully. Therefore, it is necessary to design a mechanism to improve workers' productivity with a larger prosocial reward size without deteriorating their output

quality.

In this paper, we posit that the prosocial incentives are comparable to monetary incentives in motivating workers when workers set their own goals. That is, when workers can choose their output level, their performances would increase with the prosocial reward size. We implement the prosocial incentive with a self-chosen goal scheme for several reasons. First, in previous studies, goal-setting has been used to motivate workers to improve their performance (Locke and Latham, 1990; Schroeder and Fishbach, 2015). The intuition of these studies is that goal can work as a reference point in one's decision making, leading individuals to achieve their goals without experiencing the disutility from failure[1]. Second, a few recent studies, such as Corgnet, Gómez-Miñambres, and Hernán-Gonzalez, 2015, Dalton, Jimenez, andNoussair (2015), and S. J. Goerg and Kube (2012) indicate that the information asymmetry between employees and managers would make it difficult the managers to set accurate and appropriate goals for their employees. In a real working environment, self-chosen goals might be more efficient in motivating workers than the given goals. Third, under a self-chosen goal scheme, workers retain their initial intrinsic motivation and increase their contribution quantity without sacrificing the quality, which solves reduced professionalism (Y. Li and M. S. Lee, 2013). Fourth, the self-chosen goal can also serve as a robust self-commitment tool (Klein and S. Lee, 2006; Presslee, Vance, and Webb, 2013; Brookins, Goerg, and Kube,2017). When workers' performance goals are positively related to charitable contributions, these self-commitment goals would help workers achieve their desired output levels and let a charity receive the committed donations made by the workers[2]. Such

[1]Literature shows that extrinsic motivation may crowd out intrinsic motivation in one's performance (Deci,Koestner, and Ryan, 1999, Benabou and Tirole, 2003). However, recent research has documented that the self-chosen goal can be an effective strategy to lower the salience of extrinsic motivation, emphasize a sense of achievement, which can restore workers' initial intrinsic motivation of participation (Qiao et al., 2017) and improve their performance (Corgnet, Gómez-Miñambres, and Hernán-Gonzalez, 2015). Also, individuals have sufficient information about themselves when they are setting their goals, so that they will not further undermine their motivation even when being provided with a small incentive or knowing that they are low ability (Corgnet, Gómez-Miñambres, and Hernán-Gonzalez, 2015, Corgnet, Gómez-Miñambres,and Hernán-Gonzalez, 2018).

[2]Schwartz et al. (2019) showed one reason why some literature failed to find prosocial incentives to motivate individuals better because participants could easily opt-out. Based on this finding, we can reduce the opt-

a setting would also reduce the "warm-glow giving" effect. With self-chosen goals, workers now do care about the amount of charitable contribution they commit.

To investigate the effects of self-chosen goals on workers' motivation and effort, we design an online experiment by recruiting subjects on Amazon Mechanical Turk[3]. Subjects participate in real-effort tasks, in which their performance can be associated with additional rewards that will be paid to them or donated to charities. We test our hypotheses using a 2x2 factorial design and a control group for comparison. Treatment groups vary on the type of incentives (prosocial vs. monetary) and the size of the reward (small vs. large). A total of 250 participants are randomly assigned to one of the following groups (treatments): NOINC: control group (no incentive); MONE-S: monetary incentive with small rewards; MONE-L: monetary incentive with large rewards; PROS-S: prosocial incentive with small rewards; and PROS-L: prosocial incentive with large rewards.

The experiment consists of five parts. The first part includes demographic questions and questions that help to identify subjects' commitment to a cause. This is important because previous works by Smith (2016), Carpenter and Gong (2016) found that individuals worked harder if they believed their job missions. In the second part, participants enter a practice round of real-effort-tasks (the modified encoding tasks, adopted from Erkal, Gangadharan, and Nikiforakis, 2011). In each task, participants are asked to translate letters into words according to a predetermined translation table. Such encoding tasks fit the purpose of this study since they reduce intrinsic motivation of doing the task itself, such as enjoyment or learning by doing the tasks (Charness, Gneezy, and Henderson, 2018; Corgnet, Gómez-Miñambres, and Hernán-Gonzalez, 2015; Benndorf, Rau, and Sölch, 2019).

In the third part, participants enter two rounds of encoding tasks in which they receive payments that depend on their performance. For each participant, the payment his received

out possibility by raising the cost of opt-out. That is to say, prosocial reward in our design is closely related to one's performance, so that setting a goal to oneself becomes a form of commitment between oneself and the charity, and reneging this commitment is costly to an individual.

[3]Mturk is reported to be effective in recruiting participants for online intervention research (See Cunningham, Godinho, and Kushnir, 2017).

depends on the number of tasks he finished as well as the accuracy of his translation. Then, each participant's performance is evaluated by the amount of payment he received. Each round takes 10 minutes, and participants are expected to solve as many tasks as they can within the allotted time. For each correctly solved task, participants will be paid by \$0.02. They are required to set a goal for themselves at the beginning of each round. Different from the control group (NOINC), those in MONE-S, MONE-L, PROS-S, and PROS-L are informed that the additional monetary (or prosocial reward) payment in a lump-sum form to be given to themselves (or a charity), which *only* appears in the second round (part four). This monetary or prosocial reward is a fraction of the monetary value of their goals: \$0.002 in small reward (or \$0.01 in large reward) multiples by the number of tasks to be finished (their goal). Such reward will be given to themselves or donated to a charity if goals are achieved. In the final part, participants are asked questions about their economic preferences, such as time preferences and risk preferences elicitation (Falk et al., 2015).

The first main finding of this paper is that with the same incentive type, each participant's performance improves with the reward size, or ΔMONE-S $<$ ΔMONE-L and ΔPROS-S $<$ ΔPROS-L (the difference denotes Round 2 performance minus Round 1 performance). Moreover, we find that with the same reward size, there is no significant difference in performance improvement between different types of incentives, or ΔMONE-S = ΔPROS-S and ΔMONE-L = ΔPROS-L. The second main finding is that the self-chosen goals affect participants' behaviors. Specifically, the number of one's chosen goal positively affects her performance when the goal is achieved. Moreover, participants in PROS-S and PROS-L groups are more motivated to make efforts through their chosen goals than participants in MONE-S and MONE-L groups. On average, setting a goal for 10 more tasks lead a participant to complete 1.5 more tasks in the prosocial incentive groups, or 0.5 more tasks in the monetary incentive groups. The third main finding is that being mission matched matters when participants are provided with prosocial incentives. Participants who are mission matched set higher goals and produce more output when the reward size is large, compared to their mismatched peers. This finding is consistent with our hypothesis that people are

more motivated by missions that they believed in.

The closest study to this paper is Berger, Guo, and Presslee (2018). The authors conducted a lab experiment and found that prosocial incentives led to better performance than monetary incentives (with small rewards), and a higher goal-commitment in the prosocial rewards group explains this difference. Our experimental design differs from theirs in a few settings. First of all, goals were exogenously given in their experiment, i.e., the experimenter set either a "moderate" goal (50% of participants could attain) or a "difficult" goal (20% of participants could attain) to their participants. In our experiment, a self-chosen goal is more suitable for gig workers, because it lowers the loss caused by asymmetric information (Corgnet, Gómez-Miñambres, and Hernán-Gonzalez, 2015; Dalton, Jimenez, and Noussair,2015). In addition, compared to their paper, setting a goal weighs more in our experimental setting. Indeed, a self-chosen goal relies on individuals' motivated behaviors and attitudes toward the different incentive types, reward sizes, and whether they care about charities' missions or not. Second, they recruited student subjects as participants. However, in this paper, we focus on studying the gig workers' behaviors so that recruiting Mturk workers as participants is a feasible alternative for our research purpose[4]. Third, the compensation scheme in their paper was flat wage with the prosocial reward paid if the goal was attained. In this paper, we adopt the piece-rate payment with prosocial reward to better reflect real workplace compensation.

The rest of this paper proceeds as follows. Section 1.2 introduces the theoretical framework and hypotheses. Section 1.3 presents the experimental design. The main results are exposed in Section 1.4, with the discussion of results' robustness in Section 1.5. Section 1.6 concludes.

[4]"Gig economy" has been widely recognized and becoming an essential part of organizations. Enterprising and Informal Work Activities (EIWA) Survey showed about 36 percent of respondents were undertaking informal that could be defined as "gig" works in 2015. Other studies showed that online platforms would become the primary source of providing gig works (Farrell and Greig, 2017; Abraham et al., 2017).

1.2 Model

Recent studies provide the researchers with a few mechanisms as theoretical frameworks to understand how prosocial incentives motivate individuals. One of these mechanisms is derived from the reference dependence theory (Kőszegi and Rabin, 2006; Abeler et al., 2011). A few studies such as Wu, Heath, and Larrick (2008), Gómez-Miñambres (2012), and Corgnet, Gómez-Miñambres, and Hernán-Gonzalez (2015) applied this theory into a framework of exogenously given goals boost workers' performance. These studies developed models by assuming individuals had goal-dependent preferences, and the goals in their frameworks served as reference points that separating the outcomes into the gain domain, where goals were attained, and the losses domain, where goals were not attained. Their results suggested that goal-setting was effective in motivating individuals when applying monetary rewards as incentives. Dalton, Jimenez, and Noussair (2015) extended the framework by endogenizing the goal (use the self-chosen goal) with the monetary incentive. In this paper, we extend Corgnet, Gómez-Miñambres, and Hernán-Gonzalez (2015)'s and Dalton, Jimenez, and Noussair (2015)'s model with approaches to motivate workers by prosocial incentive instead of monetary incentive.

1.2.1 Environment

Consider a framework where worker i has the goal-dependent preference. Her utility function has two components: extrinsic value and intrinsic value. The extrinsic utility function of this worker is:

$$V_E = a\theta_i q + bg\mathbb{1}_{\theta_i q \geq g} \tag{1}$$

Here, a represents the piece-rate payment. $\theta_i \in \{\theta_L, \theta_H\}$ denotes her ability level, where θ_L stands for low ability, θ_H stands for high ability, and assume $0 < \theta_L < \theta_H$. q is the labor input (effort) chosen by this worker, and g is her production goal. $\mathbb{1}$ is the indicator function, which equals 1 if the worker's output level $\theta_i q$ is greater than or equal to the setting goal of g. $b \geq 0$ is a parameter that indicates monetary rewards will be paid to the worker (or paid

to a charity on behalf of the worker) if she reaches the chosen goal. The worker (or charity) can receive larger rewards for setting a more ambitious goal, but no award will be paid if the target output level is not attained.

Besides the extrinsic utility, the worker derives intrinsic utility from achieving the chosen goal. Following Wu, Heath, and Larrick (2008) and Corgnet, Gómez-Miñambres, and Hernán-Gonzalez (2015), we define the worker's intrinsic utility function as:

$$\tilde{V}_I = \begin{cases} \mu V_I(\theta_i q - g) & \text{if } \theta_i q > g \\ -\mu\lambda V_I(g - \theta_i q) & \text{if } \theta_i q \leq g \end{cases} \tag{2}$$

Here, $\mu \geq 0$ captures the weight of the psychological component on the worker's total utility, which measures how much individuals care about achieving their goals (i.e., the intensity of goal commitment). Let $\mu \in \{\mu_m, \mu_p, \text{and } \mu_N\}$ represent the concern level of individuals with the monetary incentive, the prosocial incentive, or without incentive, respectively. $\lambda > 1$ captures the worker's degree of loss aversion. Following Wu, Heath, and Larrick (2008), we assume the functional form satisfying $V_I'(\cdot) > 0$ and $V_I''(\cdot) < 0$. These two assumptions guarantee that the value function $\tilde{V}_I$ is consistent with the properties of the value function in the Prospect Theory (Kahneman and Tversky, 1979)[5].

Assume for any $q > 0$, the cost function $c(q)$ is increasing in q and convex, that is: $c'(q) > 0$ and $c''(q) > 0$. Given the ability and goal, the worker's utility function can be written as:

$$U(q; g, \theta_i) = \begin{cases} a\theta_i q + bg + \mu V_I(\theta_i q - g) - c(q) & \text{if } \theta_i q \geq g \\ a\theta_i q - \mu\lambda V_I(g - \theta_i q) - c(q) & \text{if } \theta_i q < g \end{cases} \tag{3}$$

[5]Let $z = \theta q - g$ be the gap between output and goal. Under the assumption $\mu > 0$, $\lambda > 1$, $V_I'(\cdot) > 0$ and $V_I''(\cdot) < 0$, we can verify that: i) $V_I(0) = 0$, ii) $V_I(z) < -V_I(-z)$ for $z > 0$, iii) $V_I''(z) > 0$ for $z < 0$ and $V_I''(z) < 0$ for $z > 0$.

1.2.2 Ability is known

The timing for the worker's optimization problem is as follows. At the beginning of each period, the worker observes her ability level θ_i and sets her production goal. After choosing the goal, the worker chooses input q to maximize her overall utility $u(q; g, \theta_i)$. When the ability θ_i is known to the worker at the moment of setting a goal, it is optimal for her to set a goal such that $\theta_i q > g$ (achieving her goal). Given g and θ_i, the optimal level of q^* is given by[6]:

$$a\theta_i + \theta_i \mu V_I'(\theta_i q - g) = c'(q) \tag{4}$$

At the beginning of each period, the worker chooses optimal g^* to maximize her overall utility, given the policy function of $q(g)$:

$$\begin{aligned} \underset{g \geq 0}{max} \quad & a\theta_i q + bg + \mu V_I(\theta_i q - g) - c(q) \\ s.t. \quad & a\theta_i + \theta_i \mu V_I'(\theta_i q - g) = c'(q) \end{aligned} \tag{5}$$

Assume the function $V_I(z) = z^{\frac{1}{2}}$ and the cost function is $c(q) = \frac{q^2}{2}$, we can solve the optimal g and q for any given θ_i as:

$$\begin{aligned} q^* =& (a+b)\theta_i \\ g^* =& (a+b)\theta_i^2 - \left(\frac{\mu}{2b}\right)^2 \end{aligned} \tag{6}$$

with the parameter assumption that $\mu \leq 2\theta b\sqrt{a+b}$ is held. That is, the rewards for setting higher goal should be large enough so that the worker will not set a goal as zero. Obviously, here we find the optimal goal g^* is increasing with b and decreasing with μ:

$$\frac{\partial g}{\partial b} = \theta^2 + \frac{\mu^2}{2b^2} > 0, \quad \frac{\partial g}{\partial \mu} = -\frac{\mu}{2b^2} < 0, \quad \frac{\partial^2 g}{\partial b \partial \mu} = \frac{\mu}{b^2} > 0$$

The goal affects the worker's overall utility in two ways. On the one hand, higher goal

[6]The optimization problem is $max_q \ a\theta_i q + bg + \mu V_I(\theta_i q - g) - c(q)$. By assumption $c''(\cdot) > 0$ and $V_I''(\cdot) < 0$, so we can show that the SOC is negative: $\theta_i \mu V_I''(\theta_i q - g) - c''(q) < 0$.

brings larger monetary rewards to workers ($(a+b)\theta_i^2 > 0$). On the other hand, higher goal also reduces worker's intrinsic utility ($-\left(\frac{\mu}{2b}\right)^2 < 0$). The optimal goal g is set such that the marginal monetary utility equals the marginal intrinsic disutility. Therefore, an increase in the monetary reward size b raises the marginal monetary utility and induces the worker to set higher g. On the contrary, an increase in the psychological component μ (i.e., being more concerned about achieving the goal) makes the intrinsic disutility more negative, forcing the worker to set a lower goal.

1.2.3 Ability is unknown

Now, suppose that before the worker sets her goal, she does not observe her ability or has imperfect knowledge about her own ability. Instead, she knows the distribution of the ability level and the probability of attaining each ability level. Suppose p is the probability for attaining a higher ability level (θ_H), and 1-p is the probability for attaining a lower ability (θ_L)[7]. At the beginning of each period, the worker sets optimal g^* to maximize her expected overall utility. After the worker has set her goal, she sees the task, and her ability to do that task is revealed. Then, she chooses the labor input to maximize the utility given her chosen goal[8]:

$$\begin{aligned} \underset{g\geq 0}{max} \quad & p[a\theta_H q_H + bg + \mu V_I(\theta_H q_H - g) - c(q_H)] + (1-p)[a\theta_L q_L - \mu\lambda V_I(g - \theta_L q_L) - c(q_L)] \\ s.t. \quad & a\theta_H + \theta_H \mu V_I'(\theta_H q_H - g) = c'(q_H) \\ & a\theta_L + \theta_L \mu\lambda V_I'(g - \theta_L q_L) = c'(q_L) \end{aligned} \tag{7}$$

[7]In this paper, being a higher ability individual means one can achieve her goal. Gonzalez, Dalton, and Noussair (2020) discussed why individuals might have incomplete information about their own abilities, it could be caused by personal overconfidence or underconfidence (also see Benoit, Dubra, and Moore, 2015; Grossman and Owens, 2012; Clark and Friesen, 2009).

[8]SOC of q_H and q_L is $\theta_H^2 V_I''(\theta_H q_H - g) - c''(q_H) < 0$ and $-\theta_L^2 V_I''(g - \theta_L q_L) - c''(q_L) < 0$ respectively. The former inequality is guaranteed as $V_I''(\cdot) < 0$ and $c''(\cdot) < 0$. The second inequality is guaranteed following Lemma 1 in Wu, Heath, and Larrick (2008): $\tilde{V}_I''(\cdot) < c''(\cdot)$ by the assumption of myopic optimization.

The optimal $g^* \in (\theta_L q, \theta_H q)$ is then given by[9]:

$$bp = p\mu V_I'(\theta_H q_H - g^*) + (1-p)\mu\lambda V_I'(g^* - \theta_L q_L) \tag{8}$$

Proposition I *Under the assumption that $\mu > 0$, $V_I''(\cdot) < 0$ and $c''(\cdot) > 0$, the optimal input q_H increases with g^* when ability is high and the optimal input q_L decreases with g^* when ability is low:*

$$\frac{dq_H}{dg^*} > 0, \quad \frac{dq_L}{dg^*} < 0$$

Proof. Taking derivatives with respect to g on FOC of q_H and q_L and re-arrange it:

$$\begin{aligned}
\frac{dq_H}{dg^*} &= \frac{\mu V_I''(\theta_H q_H - g)\theta_H}{\mu V_I''(\theta_H q_H - g)\theta_H^2 - c''(q_H)} > 0 \\
\frac{dq_L}{dg^*} &= \frac{\mu V_I''(g - \theta_L q_L)\theta_L}{\mu V_I''(g - \theta_L q_L)\theta_L^2 + c''(q_L)} < 0
\end{aligned}$$

The first inequality is guaranteed as $V_I''(\cdot) < 0$ and $c''(\cdot) < 0$. The second inequality is guaranteed following Lemma 1 in Wu, Heath, and Larrick (2008): $\tilde{V}_I''(\cdot) < c''(\cdot)$ by the assumption of myopic optimization. □

This proposition implies that output and goals are complementary when the realized ability is high ($\theta = \theta_H$). This is consistent with the theoretical results in Corgnet, Gómez-Miñambres, and Hernán-Gonzalez (2015) and Wu, Heath, and Larrick (2008). The complementarity between output and goals stems from the property of diminishing sensitivity ($V''(\cdot) < 0$) in intrinsic utility.

[9]The SOC is given by: $-p\mu V_I''(\theta_H q_H - g)\left[\theta_H \frac{\partial q_H}{\partial g} - 1\right] - (1-p)\mu\lambda V_I''(g - \theta_L q_L)\left[1 - \theta_L \frac{\partial q_L}{\partial g}\right]$. By the FOC of q_H and q_L, we can rewrite SOC of g as: $-\frac{p}{\theta_H}\frac{\partial q_H}{g} + \frac{1-p}{\theta_L}\frac{\partial q_L}{q} < 0$ as $\frac{\partial q_H}{g} > 0$ and $\frac{\partial q_L}{q} < 0$ (see the proof of proposition I.)

Proposition II *Given g, the optimal q_H increases μ, the optimal q_L increase with both μ and λ.*

Proof. By FOC of q_H and q_L, we have:

$$\begin{aligned}
\frac{\partial q_H}{\partial \mu} =& \frac{V_I'(\theta_H q_H - g)\theta_H}{c''(q_H) - \mu V_I''(\theta_H q_H - g)\theta_H^2} > 0 \\
\frac{\partial q_L}{\partial \mu} =& \frac{\lambda V_I'(\theta_L q_H - g)\theta_H}{\mu V_I''(g - \theta_L q_L)\theta_L^2 + c''(q_H)} > 0 \\
\frac{\partial q_L}{\partial \lambda} =& \frac{\mu V_I'(\theta_L q_H - g)\theta_H}{\mu V_I''(g - \theta_L q_L)\theta_L^2 + c''(q_H)} > 0
\end{aligned}$$

□

Proposition III *The optimal goal g^* increases with b, decreases with μ and λ:*

$$\frac{dg^*}{db} > 0, \quad \frac{dg^*}{d\mu} < 0, \quad \frac{dg^*}{d\lambda} < 0$$

Proof. Taking derivative with respect to b on both side of g's FOC and re-arrange it:

$$\frac{dg^*}{db} = \frac{p}{-SOC} > 0$$

The inequality hold as $SOC < 0$ for the maximization problem. Taking derivative with respect to μ on both side of g's FOC and re-arrange it:

$$\frac{dg^*}{d\mu} = \frac{p}{SOC}\frac{c''(q_H)}{\theta_H}\frac{\partial q_H}{\partial \mu} + \frac{1-p}{SOC}\frac{c''(q_L)}{\theta_L}\frac{\partial q_L}{\partial \mu} < 0$$

the inequality hold as $SOC < 0$, $\frac{\partial q_H}{\partial \mu} > 0$ and $\frac{\partial q_L}{\partial \mu} > 0$. Taking derivative with respect to λ on both side of g's FOC and re-arrange it:

$$\frac{dg^*}{d\mu} = \frac{1-p}{SOC}\frac{c''(q_L)}{\theta_L}\frac{\partial q_L}{\partial \lambda} < 0$$

the inequality hold as $SOC < 0$ and $\frac{\partial q_L}{\partial \lambda} > 0$ $\square$

Proposition I and III indicate that for high-ability workers (or workers who achieve the goal), the input q_H increases with the reward size b; whereas for those low-ability workers (or workers who do not achieve the goal), the input q_L decreases with the reward size b. b's impact on input q_H and q_L are mainly through changes in g. Workers would set higher goal when facing a larger reward size b. This, by proposition I, would encourage input for high-ability workers and discourage input for low-ability workers.

Proposition I, II, and III also imply that the total effect of μ and λ on q_L is positive, in that both the direct effect and the indirect effect (the path through g) are positive. That is, when the worker cannot achieve the goal she set, her output $\theta_L q_L$ increases with her weight on its psychological component μ and the degree of loss aversion λ. Intuitively, when μ or λ higher, it is more costly for not achieving the goal. The worker would then raise the input q_L to mitigate the disutility from the psychological component and loss aversion.

However, the total effect of λ on q_H is negative (there is no direct effect, and indirect effect is negative), whereas the total effect of μ on q_H is ambiguous. With higher degree of loss aversion, the worker would set a lower goal g. As the output increases with g in high-ability workers and λ does not directly impact on q_H, her output $\theta_H q_H$ decreases with a reduced goal g. The ambiguity of μ on q_H can be see through the following equation:

$$\frac{dq_H}{d\mu} = \left[c''(q_H) - \mu V_I''(\theta_H q_H - g)\theta_H^2\right]^{-1} \left[\underbrace{V_I'(\theta_H q_H - g)\theta_H}_{\text{direct effect}>0} \underbrace{-\mu V_I''(\theta_H q_H - g)\theta_H \frac{dg}{d\mu}}_{\text{indirect effect}<0}\right]$$

Proposition II shows that the direct effect of μ on q_H is positive, whereas proposition I and III imply that the indirect effect of μ on q_H through changes in g is negative. The resulting total effect of μ on q_H is not determined. If the direct effect dominates the indirect effect, output increases with the psychological weight μ in high-ability workers. If the indirect effect dominates the direct effect, output decreases with μ among high-ability workers.

1.2.4 Mission matched and mismatched

Prosocial incentive differs from monetary incentive at its payment is contributed to charitable organizations. Therefore, if we assume a worker knows that she is contributing to a cause she believes in, then she will be more concerned about how much money she gives to this charity. In other words, we can use m to represent the mission matching level between an individual with the target charity, and the reward bg in worker's utility function is determined by m. So that we can rewrite the utility function given prosocial incentive as:

$$U(q; g, \theta_i) = \begin{cases} a\theta_i q + mbg + \mu V_I(\theta_i q - g) - c(q) & \text{if } \theta_i q \geq g \\ a\theta_i q - \mu\lambda V_I(g - \theta_i q) - c(q) & \text{if } \theta_i q < g \end{cases} \tag{9}$$

Based on the analysis in Proposition 2, we can easily deduce that given reward size b, m increases with the optimal goal g^* whether the ability is known or not, and m increases one's input q directly when the ability is known, or indirectly through a higher chosen goal when the ability is unknown but found to be high. Moreover, the effects of b on g^* and q is enhanced if a worker is being mission matched.

1.2.5 Hypotheses

Our model yields a few implications that are satisfied in both of the following conditions: an individual knows her ability, or an individual does not know her ability but she sets achievable goal for herself[10]. First of all, the model shows that the increases in the reward size b always leads to higher output θq, thereby our first hypothesis suggests:

Hypothesis 1: *Given the monetary or prosocial incentive, a worker's output increases with reward size.*

[10]We expect most workers to be satisfied with the assumption in a real working environment, so that our experiment is designed in a way of letting participants get familiar with their jobs before they setting goals that associated with additional rewards.

Moreover, output θq also can increase with worker's goal g when the ability is known, while the output of the high ability workers (those who achieve their goals) $\theta_H q_H$ increases with g when the ability is unknown. This gives us the second hypothesis:

Hypothesis 2: *Given the monetary or prosocial incentive, a worker's output increases with one's self-chosen goal conditional on achieving the goal.*

Notice that Hypotheses 1 and 2 are satisfied with both the monetary incentive and prosocial incentive. Our model suggests that there are differences between these two types of incentives. Given the prosocial incentive (not applied to the monetary incentive), if a worker agrees more with a charity's mission (i.e., when m increases), then her chosen goal g and output θq increases. This implication makes us safely draw the third hypothesis:

Hypothesis 3: *Given the prosocial incentive, being mission matched motivates a worker to set her goal higher, and then she produces more output.*

The model also discusses the determinants (psychological component, loss aversion, reward size, and being mission matched) that have a combined effect on an individual's self-chosen goal.

Corollary 1: *Higher levels of psychological component and loss aversion will decrease one's self-chosen goal, while a large reward size and being mission matched will increase one's self-chosen goal.*

Combine the suggestions in Hypothesis 2, 3, and Corollary 1, we can deduce that individuals who set higher and achievable goals indicate that they rely more on the goal-setting scheme to motivate themselves, and that their decisions of raising or lowering the goals are affected by the determinants in Corollary 1.

1.3 Experiment

1.3.1 Description of the Study

We recruited a total of 244 participants via Amazon Mechanical Turk (MTurk)[11]. They were randomly assigned to one of five groups in this experiment (2×2 factorial design with a control group). All groups used the same task description text in recruiting in order to minimize the selection bias. These groups were separated with respect to the incentive types (non-incentive vs. monetary incentive vs. prosocial incentive) and the size of the reward (small vs. large). NOINC was the control group with non-incentive; MONE-S was given monetary incentive with small rewards; MONE-L was given monetary incentive with large rewards; PROS-S was given prosocial incentive with small rewards; and PROS-L was given prosocial incentive with large rewards. There were 48, 50, 50, 48, 48 participants in these five groups completed the study, respectively. Table 1.1 shows the timeline describing the entire experiment.

The experiment consisted of four parts and took an average of 30 minutes to complete. Part I began with asking participants about their demographic information. These included questions about the participants' age, gender, income level (not forced to answer), employment status, etc.. Next, the experiment measured the matching level between individuals and the charities' missions (mission matching level). Studies such as Smith (2016), Carpenter and Gong (2016), Banuri and Keefer (2016), Burbano (2016), and Jeworrek and Mertins (2019) found a positive relationship between "mission matched" and working effort. In a recent study, Resh, Marvel, and Wen (2018) extended this empirical evidence and demonstrated the persistence of prosocial work effort among the "matched" individuals. In this experiment, we followed the methodology in Resh, Marvel, and Wen (2018) and selected five charities which were the most followed charities according to Charity Navigator in following

[11]MTurk is an internet marketplace to access thousands of worldwide survey-takers (workers) to perform requested tasks. Participants in our experiment had qualification scores 90%. To minimize the difference in transaction costs between weekdays and weekends, we distribute all batches on Amazon Mturk only on weekdays.

Table 1.1: Timeline

Part I	• Demographics (e.g. age, gender, income, employment status, etc.). • Mission match levels (charities in 5 areas): self match, society match. • Altruism level questions.
Part II	• Practice round (encryption tasks).
Part III	• Round 1: • - 10 minutes; set goal at the beginning. • - There is no reward/penalty whether achieve the goal or not. • - Piece rate: \$0.02. • - Post-task questions.
Part IV	• Round 2: • - NOINC: the same as Round 1. • - MONE-S: Round 1 + monetary w/ small reward: \$0.002 × goal (if achieved). • - MONE-L: Round 1 + monetary w/ large reward: \$0.01 × goal (if achieved). • - PROS-S: Round 1 + prosocial w/ small reward: \$0.002 × goal (if achieved). • - PROS-L: Round 1 + prosocial w/ large reward: \$0.01 × goal (if achieved). • - PROS-S, PROS-L randomly assigned 1 of 5 charities as donating target.
Part V	• Time preference & risk preference elicitation.

five areas: Doctors Without Borders (represents providing aid in medical care), The Nature Conservancy (represents protecting environment), UNICEF USA (represents help saving children's lives), DAV (Disabled American Veterans) Charitable Service Trust (represents providing direct service to veterans), and American Society for the Prevention of Cruelty to Animals (represents protecting animals).

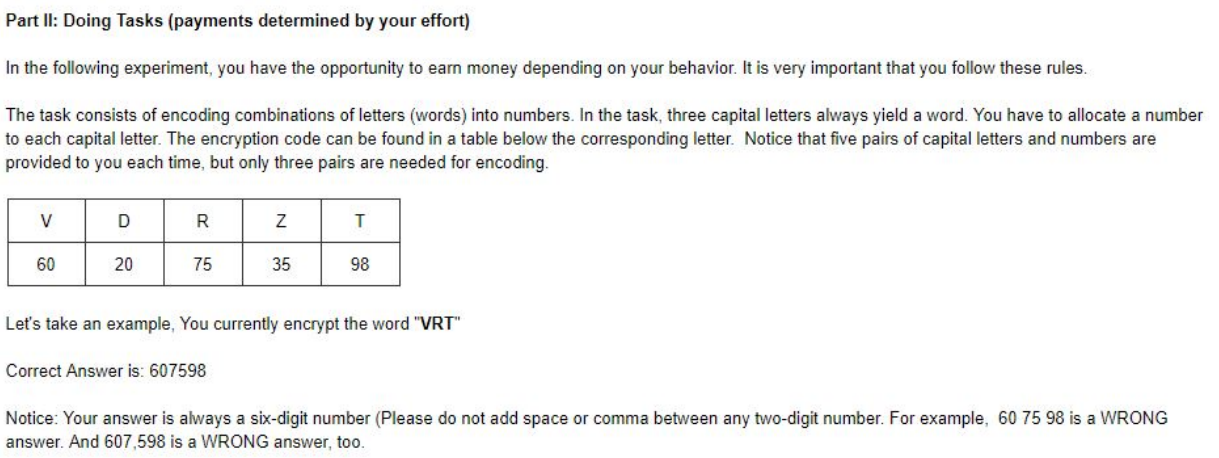
Participants were asked to read the mission statement of each charity and then asked about their attitudes towards each charity: "How much do you personally care about this issue?" and "How much do you think the society cares about this issue?" Response options ranged from 1-Not Important to 5-Very Important. The score for the first question measured how personally matched the participants were to the charity's mission (*personal matching score*, henceforth), while the score for the second question measured how important the participants thought the charity's mission was to society (*society matching score*, henceforth). At the end of part one, participants answered a series of questions eliciting their altruism level (See Falk et al., 2015). They were ten questions about self-assessment to ones' altruistic

tendencies. In our analysis in Section 1.4, we will test whether individuals' *personal matching score* affect their behaviors in the effort tasks. For example, we wondered whether the mission matched individuals (whose personal matching score for a charity is higher than the median matching score for that charity) were more motivated than their peers when they were offered the opportunity to donate to their matched charities. Moreover, in our discussion in Section 1.5, we will test whether individuals' *society matching score* and altruism levels are alternative determinants that affect ones' decision-making process.

From Part II of the experiment, participants were asked to complete real effort tasks in this experiment. First of all, they were presented with detailed instructions about the effort tasks. Participants were also asked to do a practice round, which aimed to help them fully understand the instructions and get familiar with the tasks. The effort tasks in this experiment were modified encryption tasks based on Erkal, Gangadharan, and Nikiforakis (2011). The tasks consisted of encoding combinations of the letters (words) into the numbers. Participants were provided with five capital letters in each task, and they were required to allocate a number to each capital letter using the encryption code (number) listed in a table below the corresponding letter. There were five pairs of capital letters and numbers provided in each task, but only three pairs would be used for encoding. In previous literature, encoding tasks were usually implemented in the way that letters and words had a fixed pair. Unlike these studies and in the hope of further minimizing the intrinsic motivation of doing the tasks (enjoyment or learning by doing tasks), we set the letters and numbers as randomly pairing in each task. In other words, from the perspective of the participants, each question is new for them. See Figure 1.1 for the screenshot of the instruction to the effort tasks.

Next, participants began the main part of the experiment with two rounds of encryption tasks to get compensation (Part III and Part IV). Each round took 10 minutes, and participants were expected to solve as many tasks as possible within the allotted time. They were also required to set a goal for themselves at the beginning of each round. For those in treatment groups (MONE-S, MONE-L, PROS-S, PROS-L), they were informed that the payment of additional goal reward (either monetary reward or prosocial reward) was a lump-

Figure 1.1: Effort Tasks Instruction



Part II: Doing Tasks (payments determined by your effort)

In the following experiment, you have the opportunity to earn money depending on your behavior. It is very important that you follow these rules.

The task consists of encoding combinations of letters (words) into numbers. In the task, three capital letters always yield a word. You have to allocate a number to each capital letter. The encryption code can be found in a table below the corresponding letter. Notice that five pairs of capital letters and numbers are provided to you each time, but only three pairs are needed for encoding.

V	D	R	Z	T
60	20	75	35	98

Let's take an example, You currently encrypt the word "**VRT**"

Correct Answer is: 607598

Notice: Your answer is always a six-digit number (Please do not add space or comma between any two-digit number. For example, 60 75 98 is a WRONG answer. And 607,598 is a WRONG answer, too.

sum amount, which only appeared in the second round. In Round 1, participants were paid by \$0.02 for each correctly solved task as their piece-rate compensation. Participants did set goals in Round 1, but they were also informed that there would be no goal reward or punishment whether they met their goals or not.

In Round 2, the control group (NOINC) followed the same requirements and payment method as Round 1. In contrast, treatment groups could receive goal rewards in addition to their piece-rate payment as incentives in Round 2. These rewards were paid *only if* a participant achieved his goal, and they were calculated as a fraction of the monetary value of one's goal: \$0.002 in small goal rewards groups (MONE-S and PROS-S) and \$0.01 in large goal rewards (MONE-L and PROS-L) multiples by the number of tasks they set to finish (their goals). In this paper, our main interest was to compare the performance difference across rounds (Round 2 Performance - Round 1 Performance) between groups, which reflected the effectiveness of different types of incentive in motivating workers. Moreover, participants in prosocial incentive groups (PROS-S and PROS-L) were randomly assigned one of the charities in part one (the mission matched questions) as their donating target at the beginning of Round 2. At the end of both rounds, participants were asked questions about how much they cared about achieving the goals they set when they were doing the tasks (three questions in total), which measured the psychological component in our model.

We use the Principal component analysis (PCA) approach to establish an index for this concern level. We select the first component in PCA to be this index, which explains 59.92% in the variance of Round 1 questions, and 63.66% in the variance of Round 2 questions.

In Part V, participants were required to answer two series of staircase questions in time preference and risk preference (we adopted the "staircase" procedure method from Falk et al., 2015 and Takahashi et al., 2013). Individuals' loss aversion was elicited in the risk assessment questions. In each hypothetical question, they were asked to choose between gambling (with 50% chance of getting a fixed penalty and 50% chance of receiving a positive bonus. The amount of bonus changes with questions) or receive zero payment[12]. The fixed loss here equals \$0.5 is set close to participants' expected goal reward on average in the real effort tasks (calculated from the pilot data), so that their risk preference could be elicited in a similar way as they were doing the tasks.

In addition to the groups in which participants chose their own goals, we recruited another 145 participants and randomly separated them into three groups without the goal-setting scheme. These groups are NOGOAL (no incentive; no goal-setting), NG-PROS-S (small prosocial rewards; no goal-setting), and NG-PROS-L (large prosocial rewards; no goal-setting). We compared both the outputs and performance improvements of participants with and without a goal-setting scheme. We discuss our findings in Section 1.5.2.

1.3.2 Summary Statistics

The participants in this experiment have an average age of 33.92 years old (s.d. 10.09), with about 49.59% of the participants are female. Most participants have an income level between \$50,000 to \$74,999. Around 25.82% (mainly) have part-time jobs, 43.85% have full-time jobs, 12.30% are self-employed, and the rest are students, retired, etc.. For the real effort tasks,

[12]Several screenshots of the experiment (including the mission matching questions, real effort tasks' instructions, a sample of the tasks, monetary rewards, and assigning charities to prosocial incentive participants in Round 2) are presented in the Appendix section A.4. The instructions for the entire experiment can be accessed by clicking on the following link: https://www.dropbox.com/sh/b2lrz60fmh1aw2l/AABSKvJEL3v3oqDeSq0ivNpla?dl=0.

the average earnings in this experiment are \$3.31 for a participant, and the earnings on average for groups NOINC, MONE-S, MONE-L, PROS-S, and PROS-L are: \$2.99, \$3.11, \$3.63, \$3.10, and \$3.70, respectively.[13] Furthermore, the average monetary reward earned by groups MONE-S and MONE-L are \$0.11 and \$0.60, while the average amount donated to the charities by groups PROS-S and PROS-L are \$0.11 and \$0.60, too. Namely, on average participants' goal reward equals to the 4% (small rewards) and 20% (large rewards) of one's piece-rate compensation.

For the concern level of achieving one's goal (psychological component), not surprisingly we find MONE-L scores higher than MONE-S, PROS-S, and PROS-L (p-value = 0.063, 0.047 0.019) and these four treatment groups score higher than NOINC. For the loss aversion level measurement, we observe that participants are relative loss averse on average (mean of λ = 2.19, 95% CI=[1.83, 2.56]), which is consistent with findings in the literature (Takahashiet al., 2013). Furthermore, we find that participants' loss aversion levels do not significantly different across groups.

1.4 Main Results

1.4.1 Basic comparisons of performance between groups

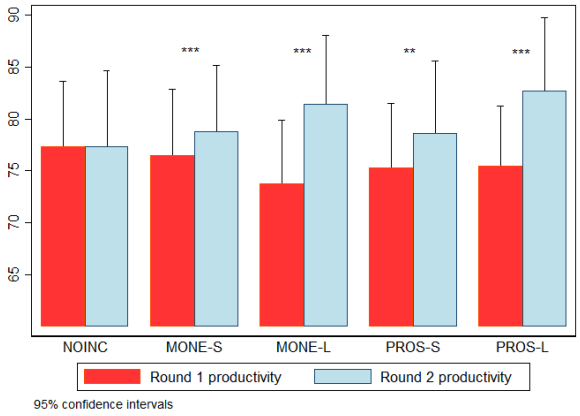
To begin with, we present the overall performance of the participants in the two-round effort tasks. While one of our experimental design's goals is to minimize the possibility of "learning by doing," we still cannot ignore the possibility that participants may improve their performance over time as they are getting more familiar with the effort task itself. This effect is supposed to exist, especially when comparing their performance (the output) across rounds. To control for this effect, we first examine the behavior of individuals in the control group, and then we compare the performance of participants in the treatment groups and the con-

[13] In our experiment, if we calculate the payment as per hour, it equals to \$6.61 if we count goal rewards compensation, or it equals to \$6.05 without taking goal rewards compensation into consideration. These two numbers are very close to the average earning for a Mturk worker (Hara et al., 2018 used 3.8 million tasks on AmazonMturk and calculated the average earning was \$6.19/h).

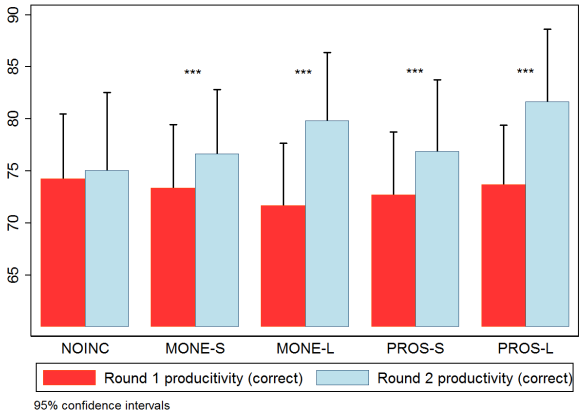
trol group. Figure 1.2 illustrates the comparisons.

FIGURE 1.2: PERFORMANCE COMPARISONS BY GROUPS

(a) Number of tasks completed



(b) Number of tasks completed correctly



This figure depicts the distribution of the participants' average effort, which is the the number of tasks completed (henceforth, *total output*) and the number of tasks completed correctly (henceforth, *actual output*) of each group[14]. Figure (a) shows that participants complete about 75 tasks on average in Round 1. While in Round 2, we do not observe the NOINC (the control group) participants' output has a statistically significant difference in their performance compared to Round 1 performance. This reflects the intrinsic motivation of doing the tasks in our experiment is limited. Nevertheless, we do see all of the treatment groups (MONE-S - PROS-L) output improves significantly. Figure (b) represents actual output across groups and shows a trend similar to figure (a). Furthermore, we show that the accuracy rates increase statistically significantly in the treatment groups, but it is not the fact in the control group (See Figure A.1 in the Appendix).

Next, we compare how much does each group improve their performance across rounds. First, we observe that there is no significant difference in the actual output between the control and treatment groups in Round 1. We then find that in Round 2, participants in

[14]In the analysis, we use an individual's actual output to refer to her output, because this output level determines whether she is rewarded or not. In the appendix, we provide the additional analysis using the output output to check for potential differences in findings.

MONE-S - PROS-L complete 3.26, 8.14, 4.15, 7.94 more tasks than their output in Round 1, respectively. These represent an increase of 4.44%, 11.36%, 5.71%, and 10.77% in ones' performance. These results are consistent with our Hypothesis 1 that participants perform better as the reward size increased, i.e., NOINC < MONE-S < MONE-L and NOINC < PROS-S < PROS-L[15]. We test the statistical significance of this finding by comparing the treatment effects between groups using a difference-in-difference approach. Our empirical specification is as follows:

$$Y_{it} = \beta_0 + \beta_1 treatment_i + \beta_2 Round_t + \beta_3 treatment_i \times Round_t + \gamma Controls_i + \epsilon_{it}$$

where Y_{it} denotes the actual output of individual i in Round t, $treatment_i$ represents a dummy variable stands for whether a participant is assigned to control group (= 0) or one of the treatment groups (= 1), and coefficient β_3 denotes the estimates to the treatment effect.

In Table 1.2, we compare the treatment effects between each treatment group (the groups with incentives) with that of the control group (the group without incentives). We find that participants in all treatment groups have significantly greater treatment effects compared to NOINC[16][17]. Our findings remain significant if we include control variables. Moreover, if we select the number of tasks completed regardless of the correctness (the total output) as the dependent variable, we obtain similar findings (see Table A.1 in the Appendix).

We compare groups with different levels of reward size under the same type of incentives in Table 1.3. Our model predicts that output increases with reward sizes given the same incentive types. Columns (1) - (2) present the comparisons within the monetary incentive groups, and indicate that MONE-L (on average) improves by 4.80 more tasks than MONE-S. While columns (3) - (4) present the comparisons within the prosocial incentive groups,

[15] Kruskal-Wallis tests suggest statistically significant differences in the median output improvement between these two sets of comparisons (p-value = 0.000; p-value = 0.017).

[16] We report the one-sided p-value for the tests according to our hypotheses.

[17] All standard errors of the parameters in this paper are obtained by bootstrap with 1,000 draws.

Table 1.2: Treatment effects between groups (actual output)

	NOINC vs. MONE-S		NOINC vs. MONE-L		NOINC vs. PROS-S		NOINC vs. PROS-L	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Complete	Complete	Complete	Complete	Complete	Complete	Complete	Complete
Treatment	−0.891	−0.307	−2.611	−0.272	−1.563	0.777	−0.563	3.717
	(4.310)	(4.652)	(4.317)	(5.462)	(4.146)	(4.933)	(4.141)	(4.299)
Round2	0.771	0.771	0.771	0.771	0.771	0.771	0.771	0.771
	(1.575)	(1.575)	(1.575)	(1.651)	(1.545)	(1.601)	(1.545)	(1.545)
Treatment × Round2	2.489*	2.489*	7.369***	7.290***	3.375*	4.012**	7.167***	7.167***
	(1.834)	(1.834)	(1.892)	(1.996)	(2.095)	(2.154)	(2.167)	(2.167)
Controls	N	Y	N	Y	N	Y	N	Y
N. Obs	196	196	196	194	192	188	192	192
R-Square	0.003	0.040	0.017	0.043	0.004	0.066	0.020	0.151

Notes: * p < 0.1; ** p < 0.05; *** p < 0.01. Dependent variable is the number of task completed correctly. Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between NOINC (control group) and MONE-S (small reward with monetary incentive). Columns (3) and (4) are comparisons between NOINC and MONE-L (large reward with monetary incentive). Columns (5) and (6) are comparisons between NOINC and PROS-S (small reward with prosocial incentive). Columns (7) and (8) are comparisons between NOINC and PROS-L (large reward with prosocial incentive). Standard errors cluster at the individual level.

and indicate that PROS-L (on average) improves by 3.16 more tasks than PROS-S. These differences are still significant at 1% level and 10% level after including the control variables. Our findings suggest that participants make higher levels of effort when provided with large reward sizes, regardless of their incentive types (monetary or prosocial).

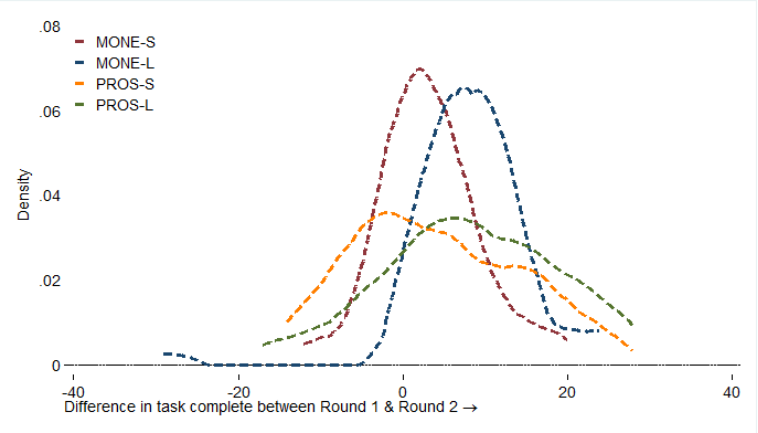
Table 1.3: Treatment effects comparisons by reward sizes (actual output)

	MONE-S vs. MONE-L		PROS-S vs. PROS-L	
	(1)	(2)	(3)	(4)
	Complete	Complete	Complete	Complete
MONE-L	−1.720	−0.469		
	(4.295)	(4.690)		
PROS-L			1.000	−1.276
			(4.136)	(5.125)
Round2	3.260***	3.260***	4.146***	4.783***
	(0.846)	(0.862)	(1.406)	(1.418)
MONE-L × Round2	4.880***	4.801***		
	(1.384)	(1.451)		
PROS-L × Round2			3.792**	3.155*
			(2.142)	(2.163)
Controls	N	Y	N	Y
N. Obs	200	198	192	188
R-Square	0.020	0.070	0.025	0.115

Notes: * p < 0.1; ** p < 0.05; *** p < 0.01. Dependent variable is the number of task completed correctly. Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between MONE-S (small reward monetary incentive) and MONE-L (large reward monetary incentive). Columns (3) and (4) are comparisons between PROS-S (small reward prosocial incentive) and PROS-L (large reward prosocial incentive). Standard errors cluster at the individual level.

In Figure 1.3, we make graphical comparisons by plotting the frequency of the distribution of changes in the output by reward sizes. Interestingly, we observe that the treatment effects distributions are similar only when incentive types are the same. Also, when comparing to the performance improvement in the prosocial incentive groups, the monetary incentive groups have higher kurtosis (monetary: 7.08 vs. prosocial: 2.27) as well as in a higher percentage of participants that obtain positive improvements (p-value = 0.033).

FIGURE 1.3: DENSITY PLOT OF PERFORMANCE IMPROVEMENT ACROSS GROUPS



Similar to the above analysis, we make another comparison between groups of the same reward sizes but different incentive types. As expected, we found no significant difference in performance improvement between the different incentive groups, regardless of their reward sizes. These effects are insignificant, whether we include control variables or not. Table 1.4 presents these results.

In summary, based on the above findings, the following relationship holds for treatment effects between groups: NOINC < MONE-S = PROS-S < MONE-L = PROS-L[18]. Our results suggest that it is possible to motivate workers to the same extent using either monetary incentive or prosocial incentives. And using larger reward size as incentives can motivate workers to a greater extent. These empirical findings are consistent with our model's pre-

[18]We do similar tests as Table 1.3 and Table 1.4 in the Appendix Table A.2 and Table A.3. The only changes in the latter two tables is that we use participants' total output as the dependent variables. We find that the relationship for treatment effects between groups still holds in Table A.2 and Table A.3.

Table 1.4: Treatment effects comparisons by incentive types (actual output)

	MONE-S vs. PROS-S		MONE-L vs. PROS-L	
	(1)	(2)	(3)	(4)
	Complete	Complete	Complete	Complete
PROS-S	−0.672	0.504		
	(3.993)	(5.076)		
PROS-L			2.048	0.139
			(4.055)	(4.408)
Round2	3.260***	3.260***	8.140***	8.061***
	(0.842)	(0.880)	(1.087)	(1.122)
PROS-S × Round2	0.886	1.523		
	(1.739)	(1.674)		
PROS-L × Round2			−0.203	−0.124
			(1.875)	(1.904)
Controls	N	Y	N	Y
N. Obs	196	192	196	194
R-Square	0.007	0.055	0.035	0.097

Notes: $*$ p < 0.1; $**$ p < 0.05; $***$ p < 0.01. Dependent variable is the number of task completed correctly (actuall output). Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between MONE-S (small reward monetary incentive) and PROS-S (small reward prosocial incentive). Columns (3) and (4) are comparisons between MONE-L (large reward monetary incentive) and PROS-L (large reward prosocial incentive). Standard errors cluster at the individual level.

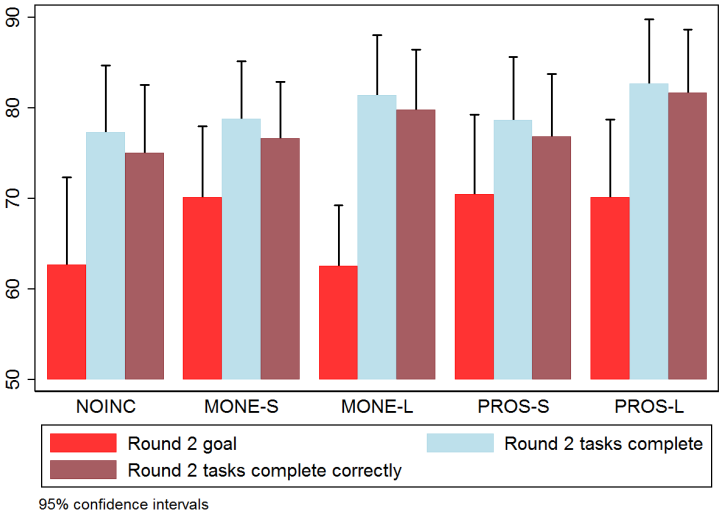
diction in Hypothesis 1.

1.4.2 Participants' goal-setting and performance improvement

Our next step is to investigate how the goal-setting scheme affects participants' behaviors. Specifically, we test whether the goal-setting scheme can explain the observed performance improvement. We begin with a descriptively summary of the average number of goals set by the participants, the number of tasks completed (total output), and the number of tasks completed correctly (actual output) in Figure 1.4. The average value of goals in Round 2 are 62.71, 70.74, 62.54, 70.48, 70.17 from group NOINC, MONE-S, MONE-L, PROS-S, and PROS-L, respectively. Moreover, the difference between the participants' output (both total output and actual output) and their self-chosen goal is greater in MONE-L than in the other groups. On average, the difference between actual output and Round 2 self-chosen goal in MONE-L is 17.26 tasks (or 18.88 tasks with respect to the total output).

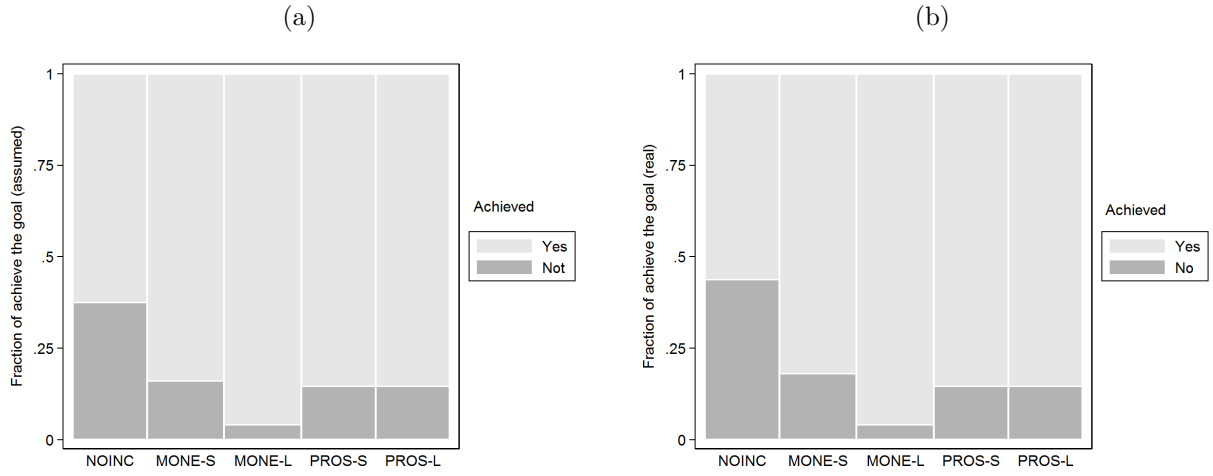
In Figure 1.5 we report each group's achievement rate, which includes both the total

Figure 1.4: Round 2 goal-setting and output



achievement rate and the actual achievement rate. The total achievement rates across groups (begin from NOINC) are 62.50%, 84.00%, 96.00%, 85.42%, 85.42%, while the actual achievement rates are 56.25%, 82.00%, 96.00%, 85.42%, 85.42%[19].

Figure 1.5: Goal achievement rate



According to Corollary 1, providing large reward size or being mission matched associated with a higher self-chosen goal, while a higher concern level about achieving one's goal (i.e.,

[19]There are only three participants in NOINC and one in MONE-S who have reached their set number of goals, but they do not achieve their targets if only the correct answers are counted as (actual) output.

a larger psychological component) or a higher level of loss aversion associated with a lower self-chosen goal. Table 1.5 displays evidence that supports these predictions. In particular, it shows that higher concern level and higher loss aversion level mainly contribute to the lower self-chosen goals in MONE-L (see columns 1 and 3, also recall that MONE-L scores higher in concern level than the other groups), while being mission matched is the main source that motivates participants with the prosocial incentive to set higher goals (see column 2). However, the effect of reward size on goal-setting is limited (see columns 1 and 2), which is supported by our experimental environment where workers are familiar with their jobs. That is, an additional reward size can directly motivate participants to make more effort, rather than necessarily updating their goal-setting strategies for indirect motivation.

Table 1.5: Determinants' effects on Goal-Setting

	MONE-S vs. MONE-L	PROS-S vs. PROS-L	NOINC vs. Treatment Groups
	(1)	(2)	(3)
	Round 2 Goal	Round 2 Goal	Round 2 Goal
Round 1 goal	0.105*	0.199**	0.185***
	(0.058)	(0.084)	(0.043)
Round 1 achieve	7.741	16.087	8.022
	(7.352)	(10.025)	(5.233)
Concern level	−8.372***	−5.831**	−5.195***
	(1.933)	(2.638)	(1.571)
Loss Aversion	−0.552*	−0.177	−0.462**
	(0.354)	(0.295)	(0.210)
Matched		17.518***	
		(5.483)	
MONE-S			9.334**
			(5.560)
MONE-L	−5.017		5.058
	(5.611)		(5.675)
PROS-S			9.207*
			(6.230)
PROS-L		−1.287	9.151**
		(6.100)	(5.494)
Controls	Y	Y	Y
N. Obs	99	94	240
R-Square	0.291	0.373	0.236

Notes: $*$ p < 0.1; $**$ p < 0.05; $***$ p < 0.01. Dependent variable is the number of goal set in Round 2. Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table displays how Round 2 goals and its interaction with the actual goal-achievement affect participants' actual output improvement.

Next, we run regressions with performance improvement on the Round 2 goal and its interaction with the goal-achievement dummy in Table 1.6. In the control variables, we

include one's concern level about achieving the goal and individual's loss aversion level in addition to the demographic variables. However, they do not provide a statistically significant impact on any group's performance, except for more loss averse decreases one's output improvement in PROS-S. Consistent with Hypothesis 2, Table 1.6 shows that for those achieve their goals, the total effort input is positively associated with their goal setting. In particular, in columns (1) - (5), those who achieve their goals increase their output by approximately 1.55, 0.24, 0.74, 1.41, and 1.63 tasks if they set 10 more goals. This reflects that participants in prosocial incentive groups are more motivated to make effort by the goal-setting scheme than monetary incentive groups. However, in column (3), we find that the higher Round 2 goal indeed improves individuals' output, but the effect of its interaction with goal-achievement on output improvement is negative. From our observation in groups' in goal-setting, we infer that this finding is caused by individuals in MONE-L set too low goals for themselves, leading to a considerable difference between the output and one's self-chosen goal, and thus the performance improvement does not depend on whether the goal is achieved or not[20].

Table 1.6: Goal-Setting's effect on Performance Improvement

	NOINC	MONE-S	MONE-L	PROS-S	PROS-L
	(1)	(2)	(3)	(4)	(5)
	ΔComplete	ΔComplete	ΔComplete	ΔComplete	ΔComplete
Round2 Goal	0.042	−0.022	0.257*	0.026	−0.046
	(0.076)	(0.038)	(0.132)	(0.086)	(0.085)
Round2 Goal × Achieve	0.113**	0.046*	−0.183**	0.115*	0.209***
	(0.063)	(0.029)	(0.104)	(0.082)	(0.074)
Setting 10 more goals (# of extra tasks completed)	1.554	0.244	0.738	1.418	1.628
Controls	Y	Y	Y	Y	Y
N. Obs	47	50	49	46	48
R-Square	0.205	0.344	0.416	0.326	0.452

Notes: * p < 0.1; ** p < 0.05; *** p < 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table displays how Round 2 goals and its interaction with the actual goal-achievement affect participants' actual output improvement.

[20]We regress the total output improvement on Round 2 goals and its interaction with the total goal-achievement in the Appendix Table A.4, and we do not see much difference from our results here.

To summarize, we find evidence supporting our Hypothesis 2 that when a participant achieves her goal, her output increases with the goal she chooses. Comparing the prosocial incentive groups to the monetary incentive groups, we find that the former have lower concern levels, which result in their higher self-chosen goals and eventually more output through the goal-setting. Moreover, knowing that treatment effects do not significantly differ between incentive types given the same reward size, we deduce that individuals with prosocial incentives rely more on intrinsic motivation (goal-setting scheme). In contrast, individuals with monetary incentives rely more on extrinsic motivation (receiving extra compensation) to improve their performance. Finally, we notice that self-chosen goals partially explains the difference between PROS-S and PROS-L[21], and we wonder if there is another determinant that moderates the effect of reward size on the output via one's self-chosen goal.

1.4.3 Prosocial incentive and mission matching

Our models show that matching with a charity's mission can improve workers in their self-chosen goals and performance. To recall, in our experiment, the prosocial incentive groups PROS-S and PROS-L face the mission matching related questions, and each participant in these two groups is randomly assigned a charity in Round 2 as her donating target. We define a participant is mission matched to her assigned charity if her personal matching score is higher than the median matching score for that charity, and the rest are mission mismatched. Thus, 13 out of 48 in PROS-S, and 15 out of 48 in PROS-L are defined as being mission matched in our experiment.

To see whether being mission matched affects PROS-S and PROS-L differently, we first do comparisons on the mission matched rate, Round 2 goal, and goal-achievement between two groups, and we find no significant difference (p-value = 0.657, 0.959, 1.000 for the compar-

[21] Using estimated marginal effects in Table 1.6, our calculation shows that when achieving ones' goals, the self-chosen goal scheme explains 78.48% of the difference in performance improvement between MONE-S and MONE-L, and it explains 35.15% between PROS-S and PROS-L.

isons between groups in the these three areas)[22]. Next, we perform the Mann–Whitney tests to see whether the difference between groups depends on participants are mission matched or not. We find participants in PROS-L set significantly higher goals (p-value = 0.029), and finish more tasks (p-value = 0.008) than PROS-S if they are mission matched. However, these differences no longer exist for mismatched participants. In addition, we do not find that goal-achievement is significantly different between the two groups, whether or not participants are mission matched.

Similarly, we make comparisons between the mission matched and the mission mismatched participants within each group. We find participants in PROS-L set higher goals (p-value = 0.000) and finish more tasks (p-value = 0.000) if they are mission matched, but there is no similar significant difference in PROS-S. To summarize, consistent with our model's predictions, mission matched participants are more likely to set challenging but achievable goals for themselves, and put in more effort to improve their work in Round 2. These effects are enhanced when they are provided with large rewards (in PROS-L). These findings are consistent with the Hypothesis 3 in our model.

Based on our observation in the above descriptive analysis, we add the mission matched dummy in addition to the goal-setting scheme in Table 1.7. Columns (1) and (2) confirm that the mission matched dummy positively contributes to participants' performance improvement, though the effects in PROS-S is positive but not statistically significant. To be specific, being mission matched leads to one's output improve by 0.97 and 7.42 more tasks in PROS-S and PROS-L. Moreover, controlling the mission matched dummy, we show that Round 2 goal still increases one's output improvement when the goal is achieved. We pool the prosocial incentive groups in column (3). In the meantime, we include the interaction term between PROS-L and the mission matched dummy, and show that the large reward positively moderates the effect of being mission matched on one's output improvement[23]. For

[22]See graphic comparisons in Figure A.2, Figure A.3, and Figure A.4 in the Appendix.

[23]The interaction term remains statically significant if we use the total output improvement and total goal-achievement dummy as the dependent variable and independent variable, see Table A.5.

example, an individual who achieves her goal of finishing 100 tasks can improve her output by 10.53 more tasks in Round 2 (about 14.38 percentage points improvement). Moreover, if the individual is mission matched and provided with large rewards, she improves her output by 6.16 more tasks in Round 2 (about 8.41 percentage points improvement).

Table 1.7: Goal-setting's effect on performance improvement with mission matched or mismatched

	PROS-S	PROS-L	PROS-S & PROS-L
	(1)	(2)	(3)
	ΔComplete	ΔComplete	ΔComplete
Round2 Goal	0.025	−0.107	−0.006
	(0.086)	(0.096)	(0.040)
Round2 Goal × Achieve	0.113*	0.197***	0.111***
	(0.083)	(0.079)	(0.036)
Matched	0.970	7.418**	0.903
	(3.711)	(4.065)	(3.159)
PROS-L			0.183
			(2.411)
PROS-L × Matched			6.158*
			(4.292)
Controls	Y	Y	Y
N. Obs	46	48	94
R-Square	0.327	0.516	0.363

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table displays how being mission matched and its interaction with the large reward size affect participants' actual output improvement. Standard errors cluster at the individual level.

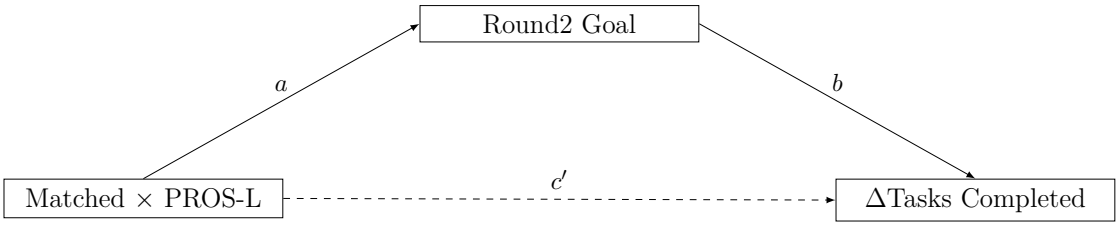
To recall, as suggested in our model, both reward size and being mission matched can affect one's output directly and indirectly (through the self-chosen goal). So that we now estimate and compare their direct effect and the indirect effect. In other words, we want to know whether the Round 2 goal plays a mediating role in this mechanism when the goal is achieved (both our model and empirical findings in Section 1.4.2 suggest that Round 2 goals significantly affect the output only if individuals achieve their goals). We find the following path has the significant effect if goals are achieved:

$$\text{Matched} \times \text{PROS-L} \xrightarrow{+} \text{Round2 Goal} \xrightarrow{+} \Delta \text{ Tasks Completed}$$

Figure 1.6 represents the framework for this path (mediation analysis). In the DAG, path a (the solid line) represents the effect of the independent variables (interaction between

PROS-L and being mission matched) on the mediator (Round 2 goal). Path b represents the mediator's effect on the dependent variable (output improvement or ΔTasks completed), whereas path c' represents the direct effect of the independent variables on the dependent variable controlling for the Round 2 goal.

FIGURE 1.6: MEDIATION ANALYSIS OF THE MECHANISM



We present the regression results of the path in Table 1.8. Column (1) analyzes path a, and shows that those who are both given large rewards and being mission matched set significantly higher Round 2 goals. Column (2) analyzes path b and path c', showing that Round 2 goals increase individuals' output improvement, after controlling the effects from PROS-L, being mission matched, and their interaction. Combining the effects in these paths, we find that the interaction has an indirect effect of $ab = 2.42$ (Sobel test p-value $= 0.071$) on one's performance improvement, which explains 36.51% of the total effect. To be specific, when an individuals is mission matched and given large reward size with prosocial incentive, she increases about 2.42 more tasks in her performance improvement through the higher self-chosen goal. Moreover, direct effect (path c') on the performance improvement equals 4.20, which is not statistically significant[24]. The intuition here is, those who have a chance to give a relatively large amount in the cause they believe in care more about how much they can contribute to the charity, and these individuals motivate themselves to improve their performance by setting higher goals. Our results are consistent with previous literature's finding that goals can mediate the effect of incentives on workers' performance (Locke and Latham, 2002; Presslee, Vance, and Webb, 2013).

[24]In Table A.6, we use total output improvement check the significance of the paths. Results are consistent with our findings here.

Table 1.8: Mediation Analysis of the Mechanism

	(1) Round2 Goal	(2) ΔComplete
Round2 Goal	NA	0.103***
		(0.042)
Matched	10.114	2.455
	(8.925)	(3.519)
PROS-L	−4.436	2.804
	(9.251)	(2.658)
PROS-L × Matched	23.435**	4.202
	(12.950)	(4.499)
Controls	Y	Y
N. Obs	80	80
R-Square	0.265	0.340

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. All columns use OLS regression. This table displays the effect of large rewards interacts with being mission matched on one's actual output improvement via the self-chosen goal. It only includes participants who achieve their goals. Standard errors cluster at the individual level.

1.4.4 Additional Findings

A. Heterogeneous effects

Tonin and Vlassopoulos (2015) and Berger, Guo, and Presslee (2018) showed that workers with relatively lower-productivity were more motivated by the monetary or prosocial rewards than their higher-productivity peers. In this paper, we obtain similar findings to these studies (See Appendix Table A.7 and Table A.8). We divide our participants, according to their productivity levels, into two groups. Namely, participants are separated by the median of their tasks finished correctly in Round 1 at 73 (mean = 73.15).[25]

More specifically, in Table A.7, we find that the treatment effects of low-productivity workers always larger than that of the high-productivity workers across groups. While Table A.8 shows that this treatment effects difference between productivity levels can be explained

[25]Table A.7 panel B and Table A.8 panel B discuss the case of participants' performance without considering the correctness (total output). We separate them by the median of tasks finished in Round 1 at 76 (mean = 75.67), the effects in these two tables are similar to our findings in the panel A parts.

by that low-productivity workers are more motivated through the goal-setting scheme. When individuals achieve their goals, setting ten more goals leads to 1.9 more tasks being completed for low-productivity workers, while for high-productivity workers it leads to 1.4 more tasks. In addition, it is more important for low-productivity workers to achieve their goals. In other words, they only significantly improve their performance if they achieve their goals. However, for high-productivity workers, even if they don't achieve their goals, they can still make some (but smaller) improvements through their goals.

B. Workers' compensation & firms' strategies

We compare participants' piece-rate earnings and rewards earned between groups in Figure A.5. We find that none of the treatment groups' (MONE-S: \$3.00, MONE-L: \$3.03, PROS-S: \$2.99, PROS-L: \$3.11) earnings is significantly higher than NOINC (\$2.99). However, when we add the piece-rate earnings and rewards together as total compensation, we find that participants in MONE-L (\$3.63) earn more than NOINC (p-value = 0.002) for themselves. Moreover, the total compensations are not significantly different between MONE-S and PROS-S (\$3.11 vs. \$3.10), or MONE-L and PROS-L (\$3.63 vs. \$3.70). Our findings here suggest that firms with different aims can have corresponding strategies in implementing the monetary incentives or prosocial incentives. Given the same amount of incentives, firms that are more concerned more about their workers' compensation can choose the monetary incentive, while firms that are more concerned about their corporate social responsibility (CSR), society's well-being, and workers' job satisfaction can choose the prosocial incentive.

1.5 Discussion

1.5.1 Goal-dependent preference vs. Goal-independent preference

Our model assumes workers have goal-dependent preferences, which indicates they are less motivated to make an extra effort after achieving their goals. In the experiment (Round 2), however, we do not know whether the participants behave strategically in the way we suggested in the model (i.e., use self-chosen goals as their reference points). Dalton, Jimenez,

and Noussair (2015) maintain that if workers do not use their goals as reference points (or they have goal-independent preferences), then their optimal output always equals to their self-chosen goals.

We test whether workers have goal-dependent preferences or goal-independent preferences in Figure A.6. We use X-axis to represent the number of Round 2 goals and Y-axis to represent the tasks completed in Round 2 (output). If workers have goal-independent preferences, then data points on the scatter plot are close to the 45-degree line. Otherwise, their outputs are slightly higher than their self-chosen goals (above the 45-degree line). This figure suggests that in our experiment, individuals' preferences are goal-dependent in all treatment groups (MONE-S - PROS-L), except for the control group (NOINC). A plausible explanation here is that although participants in NOINC set their goals in the real effort tasks, they are also told that their payments are independent of their goals. Therefore, our findings indicate that less participants in NOINC use self-chosen goals as their reference points in decision-making.

1.5.2 Self-chosen Goal vs. No Goal

To check whether the self-chosen goal scheme itself is an effective way of motivating individuals to make more effort in the real effort tasks, we add another three groups to measure and compare participants' efforts without setting their goals to their efforts with self-chosen goals. These groups are NOGOAL (no incentive; no goal-setting), NG-PROS-S (small prosocial rewards; no goal-setting), and NG-PROS-L (large prosocial rewards; no goal-setting)[26].

On average, participants in NOGOAL, NG-PROS-S, and NG-PROS-L improve their output by 1.18, 3.85, 3.83 more tasks across rounds. Using a difference-in-difference approach,

[26]For participants in NG-PROS-S and NG-PROS-L, 10% and 50% of their piece-rate payments are paid to charities as prosocial rewards, respectively (this does not affect their own compensation). On the contrary, individuals in PROS-S and PROS-L receive prosocial rewards *only if* they achieve the goals they have chosen, and the rewards are limited by the goals, rather than calculated in terms of the total correct completion of the tasks. Therefore, our measurement of the differences between groups with and without self-chosen goals represents the minimal improvement in implementing the goal-setting scheme.

we do not observe a significant difference in treatment effects between NOINC and NOGOAL (p-value = 0.911), or between PROS-S and NG-PROS-S (p-value = 0.602). However, both Round 1 and Round 2 output are significantly higher in groups that choose their own goals. Moreover, PROS-L is proved to be a more desirable treatment in motivating workers than NG-PROS-L. Participants in PROS-L produce more in both rounds, and they improve their performance far more than NG-PROS-L (p-value = 0.074).

Our findings between groups with and without goals are consistent with previous literature's suggestions that goal-setting scheme effectively motivates workers to make more effort (Brookins, Goerg, and Kube, 2017; Berger, Guo, and Presslee, 2018; Gonzalez, Dalton, and Noussair, 2020). Furthermore, our findings within groups that do not choose their goals are consistent with recent studies indication, that is, prosocial incentives (not statistically significant here) can motivate individuals improve their performance (NOGOAL vs NG-PROS-S: p-value = 0.119; NOGOAL vs. NG-PROS-L: p-value = 0.138), while individuals are not further motivated by large prosocial rewards (NG-PROS-S vs. NG-PROS-L: p-value = 0.489) (Charness, Cobo-Reyes, and Sanchez, 2016; Li, 2019; Imas, 2014; Khan, Goldsmith, and Dhar, 2020). We present the above results in Table A.9 in the Appendix.

1.5.3 Mission matching and density distribution

To recall our observation in Figure 1.3, prosocial incentive groups are flatter in density plots than monetary incentive groups. It can be important for firms to understand who are the most motivated and who are the least motivated workers given a specific incentive type, especially when it comes to the prosocial incentive groups. Inspired by our above analysis that being mission matched matters when providing participants with the prosocial rewards, we further analyze whether the distribution of performance improvement in PROS-S and PROS-L can be explained by mission matching difference between participants.

As Figure A.7 (a) shows, we separate participants in PROS-S and PROS-L by their performance improvement into quarters based on percentiles. More specifically, we study the top and bottom 25% of participants within their groups. For PROS-L, we find that the

mission matched participants are significantly more (less) likely to be the top (bottom) 25% participants than their mismatched peers (p-value = 0.019, 0.020). Similar but insignificant results are observed for PROS-S (p-value = 0.885, 0.359). Therefore, our findings indicate that mission matching indeed partly explains the flat density distribution of the prosocial incentive groups.

On the contrary, in Figure A.7 (b), we do not observe similar mission matching findings among groups without the goal-setting scheme (NG-PROS-S and NG-PROS-L). This suggests that when individuals contribute to a cause, they will select their effort strategy based on the donating targets *only* when they can choose their own goals.

1.5.4 Alternative paths in the mediation analysis?

In addition to using participants' personal mission matching level, we wonder if other determinants play a role in the path of prosocial incentive motivating individuals to make more effort through the self-chosen goals. The first possible determinant is *society's matching score* (see Section 1.3.1), which represents how important the participants thought the charity's mission was to society. The second possible determinant is individuals' *altruism level*. Using the mental accounting theory (Thaler, 1999), Presslee, Vance, and Webb (2013) maintained that individuals would consider cash earnings as in their normal account, and non-cash rewards were in another account. Thus, it is possible that altruistic individuals would be motivated by the prosocial incentive to make an effort to achieve their chosen goals, in order to benefit their non-cash account in the form of altruism.

We use two dummy variables to represent these two variables. In the experiment, participants report their scores for each charity given (5 charities in total), including their scores for charity's "matching level with the society's concern" question, and one of these 5 charities is later randomly assigned to them in Round 2 as donating target. While the "altruism level" is the average score of all ten altruism related questions in the experiment. We define a participant who think of a charity is "matched with society" and a participant is "altruistic" if their scores are greater than the median.

A. Matched with society

We present our findings in Table 1.9. Columns (1) and (2) show that neither the "matched with society" dummy nor its interaction with large reward size is affected by the goal-setting scheme to improve their performance in Round 2 (insignificant indirect effect). However, we find there are significant direct effects in the path. Column (2) shows that individuals who agree with the social importance of the charity are motivated by this mission match to make a higher effort, but their motivations are partly crowded out when they are provided with large rewards. Intuitively, individuals may regard the charities that matched with society's concerns are more favored by other people and receive more donations. Thus, when given a chance to contribute more, participants are less preferred to contribute to those well-accepted charities (from their perspectives), but to the charities favored by themselves (personal mission matched charities, see column 3). Furthermore, we use total output and total goal-achievement dummy in Table A.10, and our findings are similar to that in the Table 1.9. Hence, we believe that the heterogeneous effect from the individuals' views on charities (their donating target) cannot be ignored. Also, firms should be cautious when implementing the prosocial incentive. As shown in our experiment, providing employees with charities matched with society's concern as donating targets along with large rewards is possible to demotivate them to make effort.

B. Altruism

Similarly, we analyze to see if given the prosocial incentive, how much can one's altruism level explain her output improvement through (or not through) the self-chosen goal. To be specific, we are testing the existence of the path: $Altruistic \rightarrow Round2\ Goal \rightarrow \Delta Tasks\ Completed$. We present our findings in Table A.11, and we do not find any evidence to support that being altruistic or not matters.

Table 1.9: Mediation Analysis of the Mechanism (Other Path 1)

	(1) Round2 Goal	(2) ΔComplete	(3) ΔComplete
Round2 Goal	NA	0.115***	0.067*
		(0.040)	(0.046)
Matched with society	11.444	6.888**	9.207***
	(11.264)	(3.197)	(3.267)
PROS-L	−2.313	7.085**	8.017***
	(9.241)	(3.081)	(3.177)
PROS-L × Matched with society	4.072	−7.560**	−14.278***
	(15.310)	(4.376)	(4.657)
(Personal) Matched			4.565*
			(3.468)
PROS-L × (Personal) Matched			5.469
			(4.412)
Controls	Y	Y	Y
N. Obs	80	80	80
R-Square	0.276	0.343	0.435

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table includes participants who achieve their goals. Performance is valued by the actual output. Standard errors cluster at the individual level.

1.6 Conclusion

In this paper, we run an online experiment with Mturk workers doing real effort tasks. We implement the self-chosen goal combined with both the prosocial incentive and monetary incentive to motivate workers. We first observe that treatment effects in both two incentive types increase with reward sizes. Moreover, with the same reward sizes, the prosocial incentive has an equivalent level of motivation in performance as the monetary incentive. Next, we show our goal-setting scheme can explain the performance improvement, that is, the performance improves with the self-chosen goals when the goals are achieved. Moreover, the prosocial incentive groups are motivated by the self-chosen goals than the monetary groups. Finally, we find that being "mission matched" also contributes to prosocial incentive motivation. The effect of mission match on the performance improvement is through the self-chosen goals when participants are provided with large rewards. In conclusion, our findings support the effectiveness in the application of prosocial incentive.

Our findings contribute to the literature on the prosocial incentive. Previous research

found individuals responded to prosocial incentive only when the incentive size was extremely low, while increasing prosocial rewards would not improve workers' performance. Also, when comparing the prosocial incentive to the monetary incentive, individuals performed better under prosocial incentive when the reward was small. However, they performed better under monetary incentive when the reward comes to a moderate amount Charness, Cobo-Reyes,and Sanchez, 2016; Li, 2019; Imas, 2014; Khan, Goldsmith, and Dhar, 2020). Unlike these studies, our study introduces a goal-setting scheme combined with the application of prosocial incentive, by which we find supporting evidence in its effectiveness. We suggest prosocial incentive can be comparable to the monetary incentive in motivating workers in the application, and applying which type of incentive depends on the firms' targets.

This paper also contributes to the literature on fulfilling corporate social responsibility (CSR). Implementing CSR in the firm, such as providing employees with prosocial incentives, is believed to bring a positive reputation to the company and improve workers' motivation (Tonin and Vlassopoulos, 2015; Charness, Cobo-Reyes, and Sanchez, 2016; DellaVigna and Pope, 2018; DellaVigna, List, et al., 2020). However, some studies pointed out that merely using prosocial incentives as a tool of motivation, in the hope of increasing workers' performance or firms' profits would result in backfire and damage the firms' reputation. For example, Cassar and Meier (2017) conducted a field experiment on Mturk, and found workers react negatively if they believed that firm aimed to use prosocial incentives to make profits. They suggested that CSR should be carefully and genuinely implemented in a real working environment. Our paper shows that implementing the prosocial incentive with the simple behavioral tool, such as setting one's own goals, can retain workers' intrinsic motivation and avoid the backfire problem.

This paper has several implications. Our empirical results indicate that the prosocial incentive can motivate the same level as the monetary incentive under a real working environment by setting their own goals. Hence, we suggest firms with aims of improving workers' job satisfaction and self-fulfillment or developing their CSR can apply prosocial incentive as a tool. Moreover, heterogeneity (such as level of mission matching) also can be an essential

determinant in implementing the prosocial incentive. For example, providing with charities that are personally mission matched can motivate workers through a higher self-chosen goal, while providing with charities that are matched with society's concern may demotivate workers. Thus, firms need to figure out the optimal way of implementing different types of incentives (monetary incentive and prosocial incentive) to workers with different preferences.

2 Chapter 2: Do Pledges Lead to More Volunteering?

The effect of pledges on volunteering is ambiguous. Pledges can boost volunteering because it is easier for people to accept a request without incurring the costs of donating effort right away. However, pledges allow individuals to find more ways to excuse themselves from having to volunteer. In this paper, we study the effect of pledges on volunteering using an online experiment. We find that pledges decrease immediate rejection to volunteering requests, but pledges are often reneged later, resulting in no overall change in volunteering rates. We provide a theoretical framework for understanding the conditions under which these results emerge.

2.1 Introduction

In the last thirty years, there have been important theoretical and empirical contributions to our understanding of charitable giving. In experimental economics, in particular, the growth of research in this area has been significant following the seminal work of Andreoni 1989, 1990. More recently, researchers have focused their attention on finding ways to increase donations. Pledges, which are promises that can be recanted, have emerged as a simple and promising mechanism to incentivize potential donors to give.

Meyvis, Bennett, and Oppenheimer (2011) argue that making a pledge to donate on a future date has the effect of pushing the pain of giving away into the future. Thus, pledges make it easier for people to donate. Breman (2011) shows that introducing distance between when one says "yes" to a request and the time when one must depart from the money, increases donations. However, Fosgaard and Soetevent (2018) show that oral promises, as well as signed promises, are often reneged, rendering no changes in the overall amount donated. Comparing donations in a sample of US churches, Peifer (2010) finds that donations are not higher in churches that use pledge cards. Others find that hard pledges (e.g., pledges that are signed with the participants' and others' initials or pledges that are made public) do have a positive effect on donations (see Chou et al., 2020; Sutan et al., 2018; Cotterill et

al., 2013). Finally, Andreoni and Serra-Garcia (2019) show that when participants are given an option to make a pledge to donate on a future date in addition to donating now, overall donations do not increase. In their experiments, donations increased only when those who promised to donate were sent thank-you notes.

Do pledges affect volunteering? Unfortunately, the answer to this question is not known. Although there are a plethora of experiments on charitable monetary donations, few researchers have looked at pledges in volunteering. This is despite the fact that pledges to volunteer are ubiquitous in nonprofit organizations. Nonprofits rely on volunteers to provide high-quality services. In the US, for example, hospitals and hospices depend on volunteers to offer personalized health care at no added cost (Hotchkiss et al., 2008, 2014; Kentner et al., 2003; Wells et al., 1990). The quality of their services depends on the willingness of these volunteers to fulfill their pledges. Furthermore, volunteering is a common and important human behavior. Volunteering generates a significant amount of economic value, and enhances the overall sense of well-being (Jenkinson et al., 2013).[27]

To investigate the effect of pledges on volunteering, we conducted an experiment with Amazon Mechanical Turk workers. Participants were randomly assigned to one of four experimental treatments. In all treatments, all participants were required to fill out a survey on Day 1 and on Day 4. In addition, participants were given an option to volunteer. Volunteering in our experiment involved answering additional survey questions to benefit a charity. In the experiment, three treatments differed with respect to whether participants could volunteer on Day 1 or make a pledge to volunteer on Day 4. In treatment 1 (G1), participants were asked to reject volunteering or to volunteer on Day 1. In treatment 2 (G2), participants were asked to reject volunteering or pledge to volunteer on Day 4. In treatment 3 (G3), participants were asked to reject volunteering, volunteer on Day 1, or pledge to volunteer on Day 4. We also included a monetary treatment (G2M) where participants could pledge to

[27]The time value of annual volunteering is estimated to be in the thousands of dollars per person, and volunteering constitutes almost 50% of the estimated value of all charitable giving. If volunteering were counted as an industrial sector of the US economy, it would represent about 1% of annual GDP.

donate money on Day 4, which compares effort donation pledges (G2) to monetary donation pledges (G2M).

When comparing behavior across our experimental treatments, we find that a larger percentage of participants would reject a volunteering request immediately when pledges are not allowed, compared to when pledges are possible as in G2 and G3. In our experiment, 69.63% of participants reject a volunteering request on Day 1, compared to around 60% when pledges to volunteer on Day 4 are allowed. However, pledges are not binding commitments and can be recanted. Despite more rejecting in G1, volunteering rates do not differ when comparing G1 (volunteer now) and G2 (pledge to volunteer). When more options are available like in G3 (volunteer now or pledge), we show that reneging rates increase. Reneging rates on pledges to volunteer on Day 4 are about 11% higher in G3 than in G2. All in all, we find that volunteering rates do not differ across the three experimental groups. We also find that the proportion of those who volunteer when only pledges are allowed (G2) is 12.15% higher than those who donate money (G2M).

When comparing behavior across our experimental treatments, we find that a larger percentage of participants would reject a volunteering request immediately when pledges are not allowed, compared to when pledges are possible as in G2 and G3. In our experiment, 69.63% of participants reject a volunteering request on Day 1, compared to around 60% when pledges to volunteer on Day 4 are allowed. However, pledges are not binding commitments and can be recanted. Despite more rejecting in G1, volunteering rates do not differ when comparing G1 (volunteer now) and G2 (pledge to volunteer). When more options are available like in G3 (volunteer now or pledge), we show that reneging rates increase. Reneging rates on pledges to volunteer on Day 4 are about 11% higher in G3 than in G2. All in all, we find that volunteering rates do not differ across the three experimental groups. We also find that the proportion of those who volunteer when only pledges are allowed (G2) is 12.15% higher than those who donate money (G2M).

To help us understand the observed experimental behaviors, we developed a simple theoretical model which describes the choices participants made in each of the four treatments.

Following the work of Rasul and Huck (2010), Damgåard and Gravert (2017), Andreoni and Serra-Garcia (2019), and Breman (2011), we incorporated altruism, warm-glow, the cost of rejecting, the cost of reneging, and participation costs into the utility functions. Our simple model provides a framework for understanding the experimental results.

The rest of the paper is organized as follows. Section 2 describes the experimental design and procedures. Section 3 presents the empirical results of the experiment. In Section 4, we introduce a theoretical framework to explain these results. In Section 5 we discuss ways in which our experimental setting can be extended.

2.2 Experimental Design and Procedures

2.2.1 Description of the Study

Our experiment consisted of three effort donation treatments and a monetary donation treatment. In all treatments, participants took part in surveys at two different times: Day 1 and Day 4. The treatments differed with respect to when participants could volunteer. We recruited a total of 649 participants via Amazon Mechanical Turk (MTurk) who were randomly assigned to one of the four treatments.[28] Participants assigned to Group 1 (G1) were given the option to volunteer for charity on Day 1 or not volunteer at all (i.e., reject the request). Those assigned to Group 2 (G2) could choose to reject a volunteering request or pledge to volunteer three days later (on Day 4). Those assigned to Group 3 (G3) were given the options to volunteer on Day 1, pledge to volunteer on Day 4, or reject a volunteering request. On Day 4, participants in G2 and G3 who pledged to volunteer in three days were given the option to confirm or renege on their pledges to volunteer. We also added a monetary donation treatment similar to G2, G2M. In this treatment, participants were asked to donate money instead of effort to benefit the charity. Table 2.1 describes the treatments.

[28] MTurk is an internet marketplace to access thousands of worldwide survey-takers (Mturk workers) to perform requested tasks. Our participants were all residents of the United States with qualification scores $\geq 80\%$.

Table 2.1: Experimental Treatments

Treatments	Time Periods: Day 1	Time Periods: Day 4
G1	Volunteer now *or* Reject	n.a.
G2	Pledge to Volunteer *or* Reject	Confirm Pledge *or* Renege
G3	Volunteer Now *or* Pledge to Volunteer *or* Reject	Confirm Pledge *or* Renege
G2M	Pledge to Donate Money *or* Reject	Confirm Pledge *or* Renege

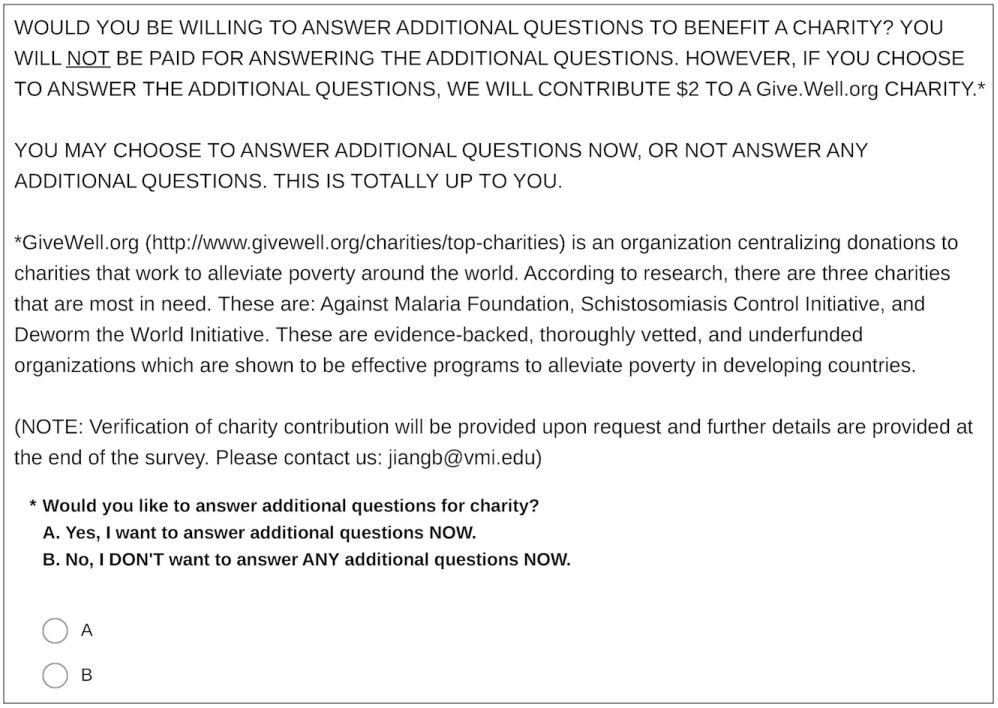
The experiment was conducted online via SurveyMonkey. Each experimental treatment consisted of several online sessions with 20-30 MTurk workers in each session. Each MTurk worker was allowed to answer only one Human Intelligence Task (HIT) on Day 1 and another on Day 4. All Day 1 questionnaires were either completed or terminated on the same day. On Day 4, at around the same time, we emailed workers who completed the surveys on Day 1 through MTurk. In the email, we included specific information about each worker's HIT that was needed to access the correct follow-up survey on Day 4. The time gap between the completion of the Day 1 questionnaire and the completion of the Day 4 questionnaire was about 72 hours (mean 71.56, S.D. 17.91).

Day 1 Survey:

On Day 1, participants answered survey questions that included demographic questions and a set of "filler" quizzes. These included questions about exercise and eating, IQ-type questions (Grosswirth and Salny, 1983), a quantitative time preference survey (Falk et al., 2016), and self-efficacy questions (Rotter, 1966). On average, the Day 1 survey took participants 11.5 minutes to complete. At the end of the survey, all participants were directed to the decision part and were asked whether they would be willing to answer additional questions

to benefit a GiveWell charity of their choice.[29] We call this additional survey the "volunteer survey". Participants in all groups were asked to confirm their initial donation decisions. A screen-shot of the donation decision for G1 on Day 1 is shown in Figure 2.1. The rest of the groups had similar choices.[30]. At the end of the survey, we asked participants to explain their choices.

FIGURE 2.1: DONATION DECISION SCREEN-SHOT FOR GROUP 1 ON DAY 1



[29] GiveWell.org (http://www.givewell.org/charities/top-charities) is an organization centralizing donations to charities that work to alleviate poverty around the world. According to research, there are three charities that are most in need: Against Malaria Foundation, Schistosomiasis Control Initiative, and Deworm the World Initiative. These are evidence-backed, thoroughly vetted, and underfunded organizations which are shown to be effective programs to alleviate poverty in developing countries.

[30] The instructions for the entire experiment can be accessed by clicking on the following link: https://www.dropbox.com/sh/ktwxyeutmjpyoqz/AADwI9p8lnMDpWwU-T7_ouela?dl=0

Volunteer Survey:

In the "volunteer survey", we asked 25 multiple-choice questions from Falk et al. (2016) as well as other "filler" questions we did not use for empirical analysis. On average, the "volunteer survey" took participants about 7 minutes to complete. To check if participants correctly estimated the time it would take them to complete the "volunteer survey", we ran a test with additional 32 MTurk workers who answered questions as in G1. We asked them two questions: *"How long (in minutes) do you anticipate it will take you to complete the additional questions to benefit a GiveWell charity?"* and *"In monetary terms, how much do you think completing the additional questions will cost you (in cents)?"* The participants anticipated that completing the "volunteer survey" would take them 6.87 minutes on average (S.D. 4.15). The estimated average time was statistically the same as how long it actually took them to complete the "volunteer survey".[31]

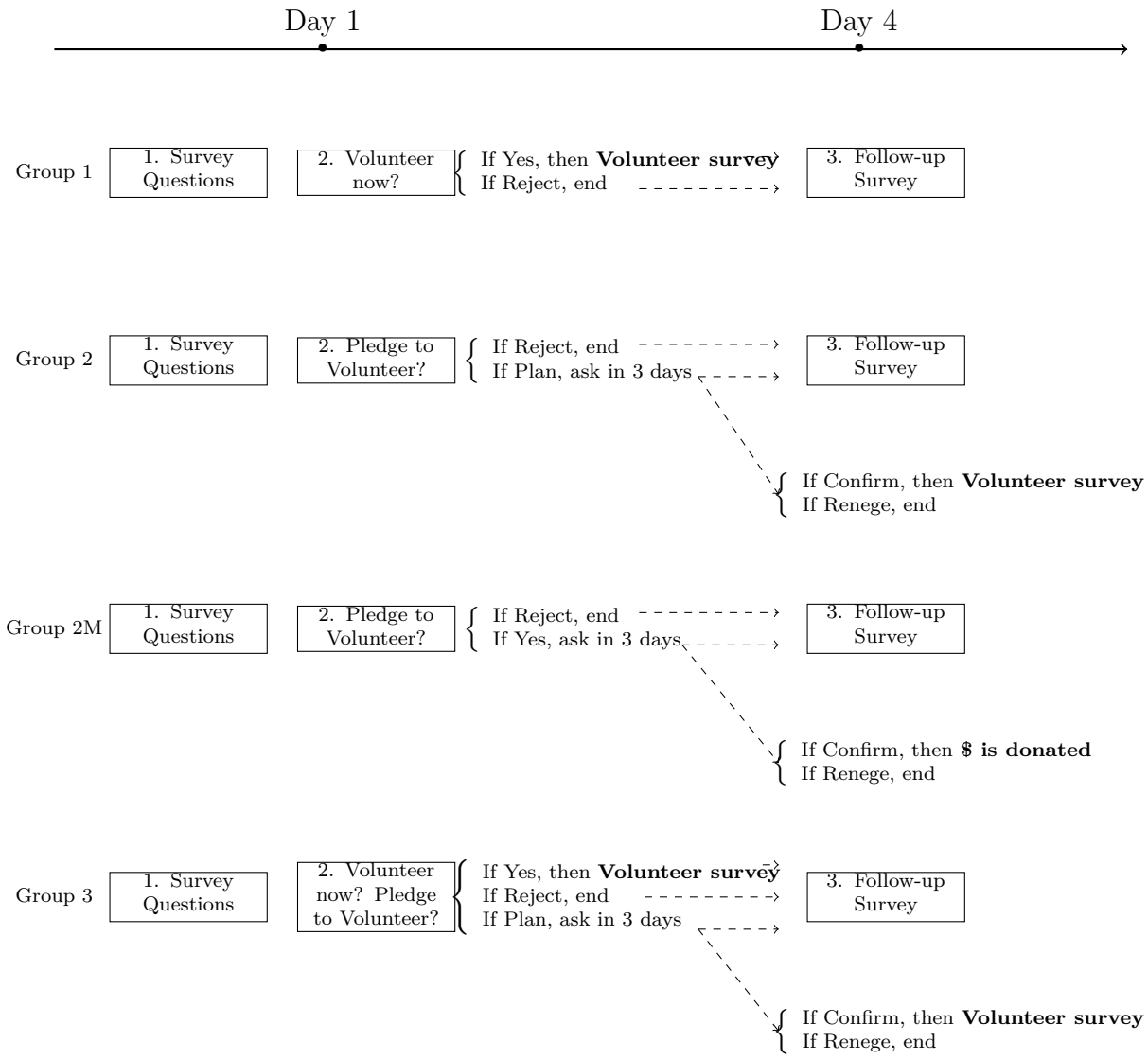
Day 4 Survey:

All participants in our experiment were required to participate in the follow-up survey three days later. This follow-up survey included a set of "filler" psychological personality questions (BFI, John and Srivastava, 1999). In addition, we asked questions to assess their altruism level and their time preferences. We included both qualitative and quantitative questions as in Falk et al., 2016. Those in G2, G2M and G3 who chose to answer the "volunteer survey" in three days were also given the option to confirm or renege on their choices on Day 4. If on Day 4, those who planned to volunteer chose to confirm, they were directed to the "volunteer survey". In contrast, if they reneged, they would skip the "volunteer survey". At the end of the entire survey, participants were asked to explain why they chose to volunteer or to renege on a plan to volunteer. Table 2.2 depicts the timeline of the experiment and surveys

[31] The higher accuracy of the estimated times is likely due to the way we designed the experiment. The option to volunteer was asked right after participants finished the first part of the survey. Lorko et al. (2019) found persisting anchoring effects when participants were asked to estimate how long they needed to complete the task.

given on each of the two days.

Table 2.2: Time-line of the Experiment



Earnings:

All participants received $2 participation fee for completing the survey on each day. If they finished both surveys, they received a total of $4 upon completion of the follow-up survey on Day 4. Those in G2M were given an additional $1 bonus. In the effort donation treatments, we contributed $2 to one of the designated GiveWell.org charities for each participant who answered the "volunteer survey". In the monetary giving treatment, G2M, participants were

given an option to donate $1 to GiveWell.org on Day 4 or reject to donate. For those who donated, we matched the $1 donation with a $1 matched gift to the charity so that GiveWell would get exactly $2. This meant that the welfare effect of donating effort and donating money was exactly the same across the G2 and G2M treatments. A fixed monetary donation of $1 for G2M "matched" the expected opportunity cost of filling out the "volunteer survey".[32]

2.2.2 Subject pool

Our subject pool consisted of 649 MTurk workers. Following recommendations from previous studies on ensuring quality responses, we recruited workers with high approval rates (see Arechar et al., 2018 and Peer et al., 2014). In our study, all recruited MTurk workers had an approval rate of at least 80%, and the vast majority of them had a 90% approval rate. In addition, we asked all of our participants to answer the time quantitative survey twice: on Day 1 and on Day 4 to test if participants paid proper attention to the tasks. We found that, on average, their elicited discount rates were consistent (Wilcoxon signed-rank test, $p = 0.460$), suggesting that participants were careful and responsible when answering questions.

The total number of MTurk workers who participated in the three effort donation treatments was 542 (272 males and 270 females); 107 workers (41 males and 66 females) participated in the monetary donation treatment. Our participants were between 20 and 73 years old with an average age of 38.73 (S.D. 10.58). About 80% of the participants were white, and two-thirds had at least some college education. See Table 2.3 for summary statistics of our participants in each treatment. We released 142, 198, and 259 HITs for G1, G2, and

[32]MTurk workers are paid by tasks completed and experienced workers keep track of the time it takes them to complete tasks. The opportunity cost of the time spent on answering the "volunteer survey" for charity equaled what workers would have made had they completed an alternative short survey, which was about $1. In a review study on MTurk workers' compensation, Hara et al. (2018) found that MTurk workers receive $6.19/h or about $0.1032 per minute. Given that participants anticipated that the "volunteer survey" was going to take less time to complete (see Table 2.2), the expected opportunity cost of volunteering was less than $1.19 (i.e., equal to the value of the average length of time on completing the surveys prior to the decision task – 11.56 minutes × $0.1032). To further validate this, we recruited an additional sample of 32 MTurk workers who were asked to estimate the expected monetary cost of finishing the "volunteer survey". The median response was $1.01.

G3, respectively. A larger number of HITs for G2 and G3 provided us with enough power to compare volunteering rates, pledge rates, and pledge reneging rates across different treatments. If a participant accepted a HIT, she instantaneously became the participant of one of our treatment groups. On Day 4, we emailed participants via MTurk and notified them to complete the follow-up HITs that linked them to the proper surveys based on their treatment groups and the choices they had made on Day 1. Of the 542 participants who completed both surveys, 135 subjects (71 males and 64 females) were in Group 1, 178 participants (85 males and 93 females) were in Group 2, and 229 participants (116 males and 113 females) were in Group 3. We released 136 HITs for G2M and 107 participants completed the HITs (65 females, 41 males, 1 other). On average, 88.30% of participants returned and completed the follow-up survey on Day 4 (completion rates were 95.07%, 89.90%, 88.42%, 78.68% in Groups 1, 2, 3, and 2M respectively).[33]

TABLE 2.3: DESCRIPTIVE STATISTICS (%)

	Average/Proportions in Sample	**Average/Proportions by Groups**			
		Group 1	**Group 2**	**Group 3**	**Group 2M**
Age (Average)	38.73	36.68	39.18	40.54	36.67
Female (=1)	51.62	47.41	52.25	49.34	60.75
White (=1)	79.04	82.22	79.78	78.60	74.77
Number of Siblings (Average)	1.98	2.01	2.02	1.98	1.88
Number of Children (Average)	0.93	0.82	1.15	0.84	0.91
Have Religion (=1)	44.68	45.93	44.38	46.29	40.19
College (=1)	65.95	65.93	61.80	69.00	66.36
Mensa Quiz ($0 \sim 5$):					
Average Score	2.60	2.56	2.49	2.66	2.78
Confidence ($0 \sim 4$):					
Average Score	2.65	2.71	2.67	2.53	2.84
Obs	649	135	178	229	107

[33]Overall, attrition was very low in our experiment because we sent three e-mail reminders to each worker. The relatively higher attrition rate for G2M was due to the fact that the HITs for this treatment were released during the months of November and December, when many people may be busy with the Holiday season in the US.

Our experimental design is similar to Andreoni and Serra-Garcia (2019) (henceforth ASG) in that they implemented a Give-Now treatment, a Pledge treatment, a Pledge-or-Give-Now treatment, and pledge treatments with Thank you letter. Participants were also sent an email reminder one day before the follow-up experimental session. However, our design differs from ASG in several important ways. First, we asked our participants to decide on whether to donate effort. Second, we used MTurk workers whereas ASG used undergraduate students. MTurk workers are more suitable participants for us to investigate effort donations. Indeed, MTurk workers are experienced online survey participants. Their compensation is directly linked to the time they spend answering surveys, so they are well aware of the opportunity cost of their time; that is, if they choose to fill out a survey for charity, they know that they will be giving up payment from an alternative HIT request (i.e., another survey). Third, our participants answered the follow-up survey on Day 4. In ASG, undergraduates completed the follow-up survey one week later. Fourth, ASG paid show-up fees of either \$15 or \$20 immediately upon completing the second laboratory session a week later. The high show-up fee helped in improving completion rates as undergraduate students have a relatively high transaction cost of showing-up to the lab for the second time.[34] In contrast to undergraduate student participants, MTurk workers' transaction costs (i.e., logging in on Mturk and answering questions) were the same on Day 1 or Day 4. The online platform enabled us to control the transaction costs. Finally, in ASG's experiment, participants first saw 5-minute presentations about a charity called *GiveDirectly* and then were asked if they would give \$5 of their participation fee to the charity. In our experiment, we did not show a video, but we provided a short explanation of the charity called *GiveWell.org* (see online instructions).

[34]The completion rate in ASG was around 90%. Meer and Rigbi (2013) illustrated that even small transaction cost would have a large effect on charitable giving.

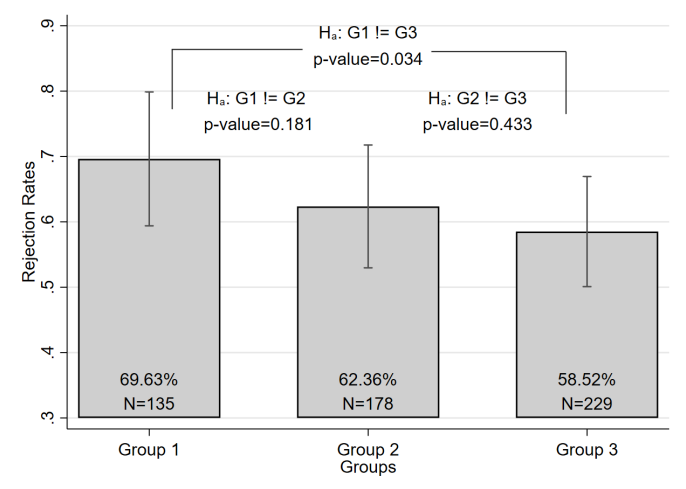
2.3 Results

As explained in Section 2.2.1, our experiment included three effort donation treatments. In G1, participants were asked to reject volunteering or to volunteer on Day 1. In G2, participants were asked to reject volunteering or pledge to volunteer on Day 4. In G3, participants were asked to reject volunteering, volunteer on Day 1, or pledge to volunteer on Day 4. In our analysis, we compare rejection rates, pledge rates, reneging rates, and overall volunteering rates across treatments. We discuss the main findings in three results below. In addition, in Section 2.4 we provide an intuitive description of a theoretical model that helps us understand these results. The details of the theoretical model can be found in the Appendix B.1.

Result 1: Rejection rates to a request to volunteer are lower when participants have the option to pledge to volunteer on a future day.

Figure 2.2 shows rejection rates on Day 1 across different treatment groups. Rejection rates in G1 are the highest, 69.63% compared to 62.36% and 58.52% in G2 and G3, respectively.

FIGURE 2.2: REJECTION RATES BY GROUPS



To investigate the effects of adding the option to pledge to donate effort on rejection rates, we compare rejection rates between G1 and G3. As shown in the Figure 2.2 rejection

rates on Day 1 are higher in G1 than in G3 (*G1 vs. G3,* $\chi^2 = 4.48$*,* $p = 0.034$). In contrast, there are no differences in rejection rates between G1 and G2. That is, pledging alone does not significantly produce less rejection compared to volunteering now. Providing people with more options (pledge and donate now) allows them to best allocate when they can volunteer. As described previously, those who expect a lower participation cost on Day 4 or are uncomfortable with the idea of saying "no" on Day 1 will rather make a pledge than reject a request.

Result 2: Reneging rates on pledges to volunteer on a future day are higher when participants have the option to volunteer now.

In both G2 and G3 participants made pledges on Day 1. On Day 4, they were asked to either confirm or renege on their pledge to volunteer. Figure 2.3 shows Day 1 pledge rates for treatments G2 and G3. In G2, 37.64% of participants pledged to volunteer whereas in G3 the rate equaled to 14.41% ($\chi^2 = 29.16$*,* $p = 0.000$). Clearly, we expected pledges to be higher in G2 than in G3 as those in G3 who wanted to volunteer could choose to do so on Day 1. However, as shown in Figure 2.4, among those who pledged to volunteer on Day 1, a larger percentage of G3 participants than G2 participants reneged on their pledge. In G2 13.43% reneged whereas in G3 24.24% reneged on their pledge ($\chi^2 = 1.83$*,* $p = 0.088$).[35] When regressing treatments on reneging rates and controlling for demographic variables we find that adding an option to volunteer now to a pledge increased the reneging rate by 11.49 percentage points ($p = 0.079$) (see Table B.3 column (2) in the Appendix).

Because in G3 participants could choose to volunteer on Day 1, pledges to volunteer on Day 4 are likely to be motivated by the avoidance of the costs of saying "no" to a request on Day 1. Andreoni and Serra-Garcia (2019) observed a similar result. ASG found that the reneging rate was 71% when students had more options (equivalent to G3) and 47%

[35]This result only includes participants who showed up on Day 4. The attrition rates on Day 4 for G2 and G3 are 5.63% and 8.33%, respectively. If we considered the no-shows as renegades, the reneging rate for G3 would be 30.56%, which is still significantly higher than the reneging rate in G2 of 18.31%.

Figure 2.3: Pledge Rates By Groups

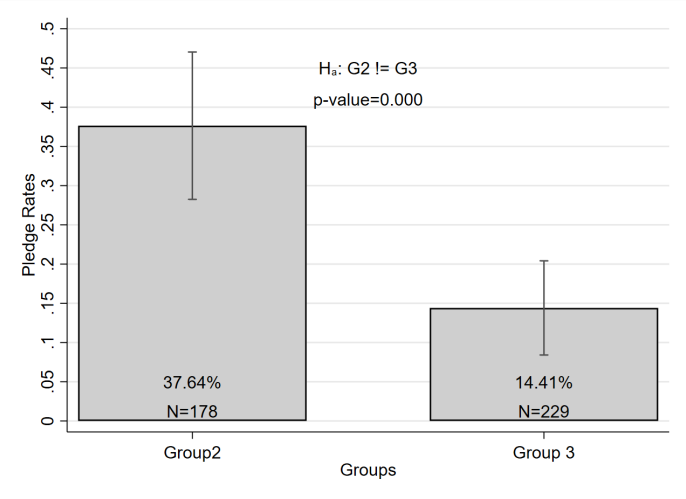
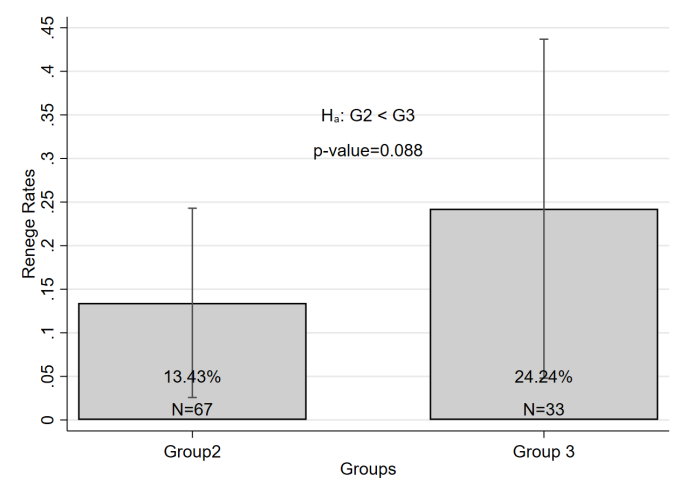


Figure 2.4: Reneging Rates By Groups



with fewer options (equivalent to G2). However, ASG showed that reneging rates were much higher than ours. Unlike our effort donation treatments with MTurk workers, ASG studied monetary donations among college students, and the duration between now and the future was longer than ours. Nevertheless, both studies come to the same conclusion that with an added option, there is a higher reneging rate.

Result 3a: There is no difference in the volunteering rates of those who are asked to volunteer now and those who are asked to make a pledge to volunteer in the future.

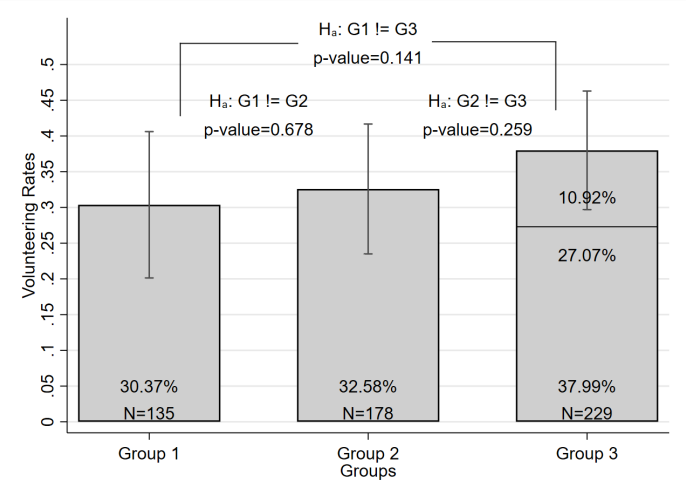
Result 3b: Adding the option to pledge to volunteer on a future day to a volunteering

request does not change volunteering rates.

Result 3c: Adding the option to volunteer now to a pledge to volunteer on a future day does not change volunteering rates.

With respect to overall volunteering, we found no statistical differences across treatments. Figure 2.5 shows the percentages of people who volunteered by treatment group. In G1, 30.34% of the participants donated effort as compared to 32.58% in G2, and 37.99% in G3 (27.07% of these participants volunteered on Day 1 and 10.92% of those who pledged to volunteer confirmed their pledges on Day 4). The volunteering rate in G3 was higher, but the difference is not statistically significant (*G1 vs. G3,* $\chi^2 = 2.16$*,* $p = 0.141$, and *G2 vs. G3,* $\chi^2 = 1.28$*,* $p = 0.259$).

FIGURE 2.5: VOLUNTEERING RATES BY GROUP



Aggregating across all treatments, 43.33% of female participants volunteered, whereas only 25.37% of male participants volunteered. This result is similar to Breman's (2011), who also found differences between men and women in how they responded to pledges. In addition, we found that the proportion of men who volunteered was higher in G3 (31.90%) than in G1 (21.13%) ($\chi^2 = 3.53$*,* $p = 0.056$). We observed no differences in the proportions

of women volunteering across treatments.[36]

Effort vs. Monetary Donations

In our experiment, we also included a treatment G2M where participants could pledge to donate money. This treatment allows us to compare effort donation pledges to monetary donation pledges.

Result 4: Effort donation rates are higher than equivalent monetary donation rates when participants are asked to make a pledge.

FIGURE 2.6: DONATION RATES BY GROUPS: MONEY VS. EFFORT

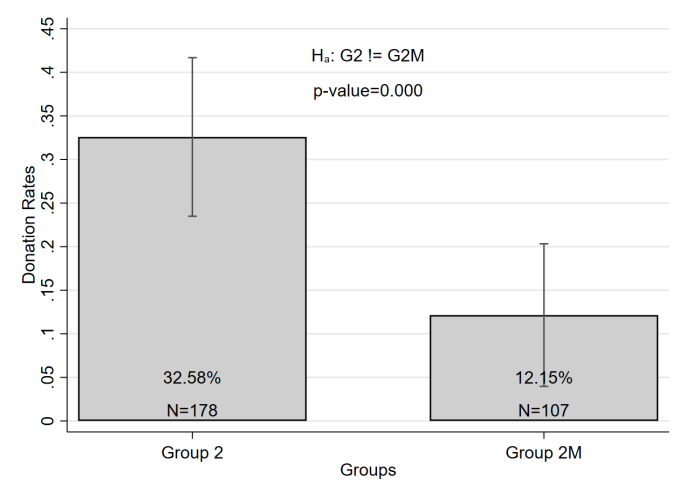


Figure 2.6 shows the rates of donation in G2 and G2M. Our data show that participants in the volunteering treatment (G2) were significantly more likely to donate than those in the monetary treatment (G2M) ($\chi^2 = 15.63,\ p < 0.001$). Existing work by Lilley and Slonim (2014), Brown et al. (2018), Tonin and Vlassopoulos (2010), and Imas (2014) supports that

[36] Although finding reasons to explain this disparity is beyond the scope of this study, we have thought about possible reasons. One of them is that the costs of rejecting effort donations increased in G3 for men, but not for women. In other words, with more options, men may have found it harder than women to say "no" to volunteering.

volunteering yields higher warm-glow utility than monetary donation, which may explain why donations are higher in G2 than in G2M.

We also observed more reneging when participants pledged to volunteer compared to monetary pledges. The reneging rate in G2 was 13.43%, which was higher than the rate of 7.14% in G2M, although the difference was not statistically significant. Higher donations and higher reneging when asked to pledge effort vs. monetary donations may sound contradictory. However, it may be relatively easier to renege a promise to volunteer time and effort than a promise to give money. When promising to donate time and effort, one can always find an excuse to renege when the future arrives by simply saying, "I really thought I could help, but it turns out I don't have time", or "Yesterday I wanted to help, but I don't feel well today", and so on. While on the other hand, excuses like these are not possible or hard to make when people promise to donate money but renege in the future. In Section 2.4 we present implications of a model that help us explain this result as well as Results 1 - 3 (see also Appendix).

Andreoni and Serra-Garcia (2019) ran a treatment similar to our G2M treatment but with student subjects. They found that the reneging rate was 47%, which was much higher than our rate of 7.14%. This difference is likely to be due to the fact that their experiment was different from ours in important ways. To begin, they used student subjects and we used MTurk workers. The average age of our subjects was 39, which is about 15 years older than the average student subject. Second, their subjects faced a higher opportunity cost of participating, because they needed to show up to the lab and participate in an experiment that lasted about one hour. Our MTurk subjects logged in from their computers and spent about 20 minutes answering questions. Third, the duration between the first survey and the second was 1 week in their experiment compared to the 3-day lag in ours. This made it harder for the students to come back to the lab. In fact, ASG endowed subjects with $6 and $20 (or $15 and $15) on Day 1 and Day 8, respectively, to encourage participation. Finally, the donation fee was $5 in their experiment, compared to our $1 match donation.

2.4 Theoretical Framework

To provide an explanation for our results, we developed a simple theoretical model that we explain in the Appendix. The model provides a framework for understanding the results of the experiment as described in the previous section. In our model, individuals compare the expected utilities of different volunteering options available to them and choose the option that yields higher expected utility. Consider a participant in G2. On Day 1, she faces two options: make a pledge to volunteer or choose not to volunteer in three days. If she pledges to volunteer on Day 1, she has the option of reneging or confirming her pledge on Day 4. She should pledge on Day 1 if and only if the expected utility of such option exceeds the utility of not volunteering. She would confirm her pledge on Day 4 if and only if the utility of confirming the pledge is greater than the utility of reneging. Following the work of Rasul and Huck (2010), Damgåard and Gravert (2017), Andreoni and Serra-Garcia (2019), and Breman (2011), we incorporate altruism, warm-glow, the cost of rejecting, the cost of reneging, and participation costs into the utility functions. So, for example, a participant in G2 who experiences high altruism (i.e., cares about the total monetary contribution to the Givewell.org charity) and high warm-glow (i.e., likes giving) will pledge and confirm her pledge if she also faces high rejection costs, hates reneging on pledges, and faces low participation costs. In the experiment, participation costs include transaction costs such as logging back in the MTurk system, as well as opportunity costs such as working in alternative MTurk HITs. If, for example, future participation costs are high and/or reneging costs are low, she may choose to pledge, but not confirm her pledge. In this situation, the option of a pledge would not necessarily lead to more volunteering.

In the process of determining the optimal choice for an individual in our model, we derived critical values that identify the conditions under which volunteering is expected. A visual depiction of these critical values and expected decisions can be found in Figure 2.7. Panels (*i* - *iii*) show the critical values and corresponding decisions on Day 1 for G1, G2, and G3, respectively. For example, the critical value C separates the decisions to reject volunteering versus volunteering on Day 1. The critical value B separates rejecting

to volunteer versus pledging to volunteer. In addition to B, panel (iii) shows critical value A which separates between pledging to volunteer and volunteering on Day 1. Our model predicts that it is always the case that $B \leq C \leq A$. The last panel (iv) shows decisions on Day 4 for individuals in G2 and G3. These individuals faced the decision to renege their pledge or confirm it. The reneging versus confirming choices on Day 4 are separated by critical value D, whereas D' represents the discounted value of D on Day 1 and is determined by reneging and participation costs (See Appendix B.1). We compare D' to the values of A, B, C to determine whether an individual would confirm pledged donation or renege on a pledge to donate. In panel (iv), the pledged donation is represented by the region between the values of B and A. However, D' could fall between B and A or be less than B or greater than A. For example, if $D' < B$, there would not be any reneging. In contrast, if $D' > A$, then everyone in G3 reneges.

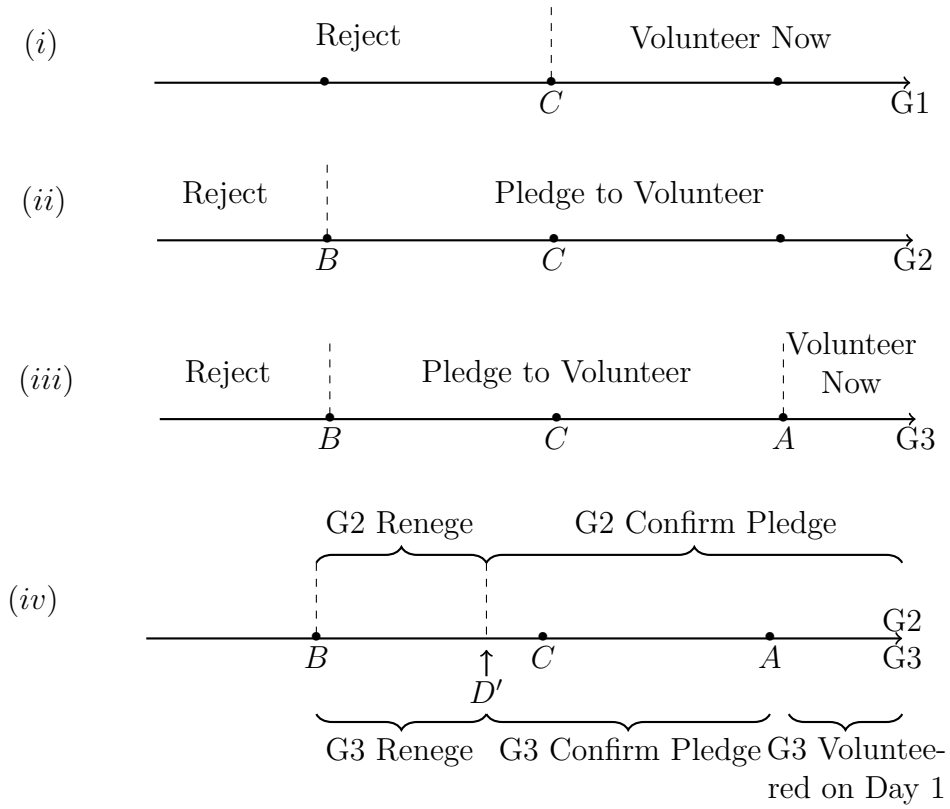
The finding that $C \leq B$, as shown in Figure 2.7, explains why rejection rates on Day 1 are higher in G1 than in G2 and G3 (see **Result 1**). Intuitively, higher rejection rates in G2 and G3 happened because participants in these treatments had more wiggle room than those in G1. That is, they had an added option to pledge and then renege later. For example, those who experienced high cost of saying "no" to a request to volunteer could easily make an empty pledge and renege later.

In the experiment, we found higher reneging rates in G3 than in G2 (**Result 2**): 13.43% in G2 and 24.24% in G3 (see Figure 2.4). As shown in panel (iv) in Figure 2.7, pledges that are reneged are determined by critical values D' and B, which are the same for both treatments. However, pledges confirmed are always smaller in G3 than in G2. This is because in G3 participants can donate on Day 1. So, the reneging rate in G2 has to be smaller than in G3.[37] Intuitively, those in G3 with high levels of altruism and warm-glow always prefer to volunteer sooner rather than later (Damgåard and Gravert, 2017)[38].

[37] In the model, the result that reneging rates are higher in G3 than in G2 holds under most conditions except for the case when no one reneges or everyone reneges. See Figure B.4 in the Appendix.

[38] See Table B.1 and B.2 in Appendix for results of a regression that shows that higher altruism level renders

FIGURE 2.7: CRITICAL VALUES
$(B \leq C \leq A)$



The model also helps us understand that when pledges are allowed as in G2 and G3, volunteering rates are affected by participation and reneging costs, which affects the value of D'. In general, the rate of volunteering in G3 should be at least as high as in G2, both of which should be strictly higher than in G1. However, when participation costs on Day 4 (i.e., the costs of logging back into MTurk and answering questions) are large and/or reneging costs are low (i.e., when $D' \geq C$), G1 can have higher volunteering rates. Moreover, when $D' \geq A$, the rate of volunteering will be higher in G3 than in G2 because all those in G3 who want to volunteer would volunteer on Day 1 (see Figure 2.1A - 2.4A in Appendix B.4 for a discussion of all possible cases).[39] In our experiment, we found no difference across

more volunteering on Day 1 and more reneging on Day 4 for those in G3 compared tho those in G2.

[39] $D' \geq A$ when reneging costs are extremely low and/or participation costs on the future date are extremely large. For example, scheduling on Christmas eve or asking participants to donate blood as in Meyer and

treatments in volunteering rates (see **Result 3a-3c**). In our experiment, we tried to make the participation costs as similar as possible on Day 1 and Day 4. However, we cannot observe reneging costs, nor can we completely control participation costs. **Results 3a-3c** suggest that D' was close to C. Our theoretical model tells us that reducing the costs of participation would push D' closer to B than to C. In this case, we would expect more volunteering when pledges are allowed. We provide some ideas on how to experimentally affect D' in the discussion (see Section 2.5). A detailed explanation of how the critical values are derived and how they pin down decisions to volunteer, pledge to volunteer, reject, and renege can be found in the Appendix Section B.1.

Effort vs. Monetary Donations

Result 4 shows that both donation rates and reneging rates were higher in effort giving than in monetary giving (donation rates were 32.58% vs. 12.15% in G2, and reneging rates were 13.43% vs. 7.15% in G2M).

The difference in donating money versus effort is the utility one experiences due to warm-glow. As shown by Lilley and Slonim (2014) and Brown et al. (2018), individuals experience higher warm-glow when donating effort as compared to donating money. Higher warm-glow in our experiment results in lower rejection rates on Day 1 or higher pledges to donate. With equal participation and reneging costs across the two Group 2 treatments, we can show that higher warm-glow effect is linked to higher overall effort donation rates in G2 compared to G2M.[40] Please see Appendix B.2 for more explanations.

With respect to reneging on a promise to donate, it is reasonable to expect that someone who dislikes refusing to donate money dislikes reneging on a promise to donate money even more. In addition, in contrast to reneging on a promise to give money, breaking the promise

Tripodi (2018).

[40] In our experiment, individuals were randomly assigned to the monetary donation and volunteering treatments. Presumably, the participation cost and reneging cost of individuals across the treatments are the same.

to donate effort by making excuses about opportunity and participation costs seems much easier. For example, one can say: "When I pledged, I thought I would have time to help, but unfortunately I don't have time right now." Indeed, 70.59% of those who reneged stated lack of time being the reason not to volunteer.[41]

2.5 Conclusion and Discussion

In this paper, we investigate experimentally the effects of pledges on volunteering. Intuitively, one would expect that when people are given more options to volunteer (i.e., volunteer on Day 1 or make a pledge to volunteer on Day 4), they would be more willing to volunteer. This is because with options, people are able to best allocate their time to help others. Indeed, a recent study by Andreoni and Serra-Garcia (2019) shows that, when people are allowed to choose when to donate money, they tend to agree to contribute at the moment they are asked, but renege on their promises later. The authors find that giving participants the option to pledge to donate does not lead to higher donations. In light of these findings, we wondered whether similar behavior would arise when people are asked to donate effort or volunteer.

To see the effect of pledges on volunteering, we designed an experiment with three treatment groups that varied with respect to whether or not participants had the option to make a pledge to volunteer, and a fourth that involved donating money. For the experiment, we recruited Amazon Mechanical Turk workers. The results show that, despite rejection rates being larger when no pledges are allowed, there are no differences in volunteering rates across treatments. We developed a simple model whereby we can identify conditions under which this result would emerge. The model tells us that, no difference in volunteering rates would happen when expected participation costs are high and expected reneging costs are low.

[41] In the model as long as reneging on a money pledge is worse than rejecting it ($|r_i^m| \leq |L_i^m|$) and finding excuses to renege on pledges to donate effort is easier than rejecting a volunteering request ($|L_i^e| \leq |r_i^e|$) and assuming that reneging effort and money is equally bad ($|L_i^e| = |L_i^m|$), the reneging rate in effort pledges is always higher than equivalent monetary pledges. Please refer to the Appendix B.2.

We find that participants renege on their pledges. Interestingly, adding the option to volunteer on Day 1 to a pledge results in higher rates of reneging on Day 4. This is because those who want to volunteer do not delay volunteering and those who pledge are probably flakes. Our model suggests that participants with high altruism or high warm-glow prefer to volunteer sooner rather than later; however, those whose costs of saying "no" are high, are driven to postpone their rejection and then renege on Day 4. Finally, when asked to pledge, volunteering rates are higher than equivalent monetary donations. This is because warm-glow effects are stronger when people donate effort (see Lilley and Slonim, 2014 and Brown et al., 2018). Compared to monetary donations, volunteering results in higher warm-glow and higher costs of saying "no".

There are a few ways one could extend our experiment. A follow-up study could be conducted by looking at the effects of participation costs (i.e., transaction costs and the opportunity costs of volunteering) on volunteering decisions. In previous studies, Knowles and Servátka (2015) found that participation costs reduced the likelihood of giving. Similarly, in two field experiments, Rasul and Huck (2010) found that a small decrease in ex-post participation costs increased the likelihood of donation by a whopping 26%. Meer and Rigbi (2013) found that even seemingly innocuous participation costs affected donations in their experiments. In an online experiment with Kiva charity recipients, presenting lenders with their own languages instead of the local language, such as Spanish and French, with a free translation option increased donations. Finally, some studies such as Damgåard and Graver (2017) and Knowles et al. (2018) suggested that shorter deadlines or without having a deadline for making donation choices increased donations, while Knowles and Servátka (2015) showed that extending deadlines had no significant effect on donations. In both our model and experiment, we assumed individuals took into account the impact of participation costs when making decisions about whether to volunteer and pledge to volunteer. To minimize the impact of large participation costs, we had individuals participate in our experiment on two fixed *weekdays* that were always three working days apart. Furthermore, participation costs in our model were assumed to be independently distributed across individuals so that

it reflected our experimental design as the costs of volunteering on a future date could be either higher or lower than the costs today. By strategically scheduling the experiment on weekends or holidays, for example, one could systematically vary participation costs of all volunteers, which may lead to different volunteering behaviors.

Other potential extensions include exogeneously decreasing the social pressure to say "no" to a donation request. For example, one could allow participants to skip the page where a decision to volunteer is posted or allow them to delete their pledges, as in Lacetera et al. (2016). Also, one could increase the costs of reneging on a promise to donate by utilizing hard pledges. For example, one can request participants to sign a non-binding agreement (a signed note or letter) or make a public statement to volunteer in the future. Andreoni and Serra-Garcia (2019) sent "thank you" notes to those who promised to donate. Chou et al. (2020) asked participants to sign their initials and others' initials. These interventions are likely to increase the psychological costs of reneging on a promise to donate, resulting in higher overall donation rates (see Bryan et al., 2010).

Finally, our findings suggest that non-profit organizations that rely on volunteers should be cautious in offering the option to pledge to volunteer in the future. Pledges may not increase volunteering, but may instead result in inefficiencies such as scheduling failures, inability to provide a promised service, and frustration caused by pledges being reneged.

3 Chapter 3: Altruistic Self-Concept Mediates the Effects of Personality Traits on Volunteering: Evidence from an Online Experiment

We study experimentally the effect of personality traits on the willingness to make and keep a promise to volunteer. In our experiment, Amazon Mechanical Turk participants were given the option to volunteer by donating time and effort to a charity. They also answered a series of questionnaires, including the Big Five personality test and attitudinal questions that we used to construct an index representing altruistic self-concept. Self-concept refers to the way we describe and evaluate ourselves. We find that altruistic self-concept mediates how personality affects volunteering decisions. In particular, agreeableness has a strong influence on the probability of making and keeping promises to volunteer *through* its effect on altruistic self-concept. Our findings have useful implications for non-profit organizations.

3.1 Introduction

There is an increasing realization among economists that personality matters. In a review of the literature, Heckman et al. (2019) show that the Big Five personality traits play an important role in predicting a wide range of life outcomes, including educational achievement, labor market outcomes, health, and criminality. In a recent study with trainee truckers data, Rustichini et al. (2016) show that personality traits have a statistically stronger predictive power than economic preferences for a diverse set of behaviors, including creditworthiness, job persistence, and vehicle accidents. Personality also predicts prosocial attitudes (Pinazo et al., 2016), prosocial choices in experimental games (Becker et al., 2012; Ben-Ner et al., 2004b; Kurzban and Houser, 2001; Volk et al., 2012), and behaviors such as donating money to charity (Yarkoni et al., 2015) and volunteering time and effort (Brown and Taylor, 2015). Building on the latter, in this paper, we investigate the effects of the Big Five personality traits on one's willingness to make and keep a promise to volunteer. In addition, we look at the effect of altruistic self-concept as a mediator of personality traits on volunteering

decisions.

The Big Five personality traits include: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Among the five traits, only agreeableness, defined as: "the tendency to act in a cooperative, unselfish manner (VandenBos, 2007, page 31)," has been consistently found to be positively correlated with prosocial behaviors.[42] Recent psychological studies argue that there may be at least one mediator between personality traits and prosocial behaviors[43]. Mediators matter because they can best identify causal pathways. Personality traits can directly influence prosociality and, indirectly, through a mediator. For example, Habashi et al. (2016) and Graziano et al. (2007) show that prosocial actions executed by agreeable individuals pass through the feelings of empathy and distress. Okun (2007) argues that the effect of personality traits on volunteering experience is mediated by contact with friends, church attendance, and (for females) membership in clubs and organizations, all of which are measures of social capital. Other authors have found that social self-efficacy (Sun et al., 2019 and Guo et al., 2018) and empathic self-efficacy (Caprara et al., 2012) mediate the effects of personality on prosocial behaviors.[44] Carlo et al. (2005) find that agreeableness and extraversion have a positive and indirect impact on the willingness to volunteer through prosocial motives. Hill (2016) finds that altruism mediates the relationship between agreeableness and organ donations. The author measures altruism through a Self-Reported Altruism Scale (SRAS), which proxies the extend to which people would behave altruistically in various situations (Rushton et al., 1981). Table C.2 in the Appendix summarizes the results of the above-mentioned studies.

In this paper, we investigate the effects of personality on the willingness to make and

[42] With respect to the other personality traits, a few papers show a positive relationship between prosociality and openness and a negative one with conscientiousness (see Table C.1 in the Appendix for a summary of highly cited studies on this topic).

[43] A mediator "transmits the effect of an independent variable on a dependent variable" (MacKinnon et al., 2007, page 1), and mediation analysis is a statistical measure estimating how much of the effect works through the mediator (Lee et al., 2019).

[44] The former relates the ability to engage in social interactions and the latter relates to a person's own perceived ability to care about others.

keep a promise to volunteer through self-concept as a mediator. Self-concept refers to how one describes and evaluates oneself, and it includes "psychological and physical characteristics, qualities, skills, roles and so forth (VandenBos, 2007, page 828)."[45] Self-concept is an adequate candidate mediator for personality, because personality traits influence self-concept and one's self-concept affects decisions and behaviors. For example, Bekkers (2006) shows that agreeableness positively develops one's altruistic self-concept, which contributes to prosocial behaviors such as blood donation. Ward and King (2018) show that agreeableness and conscientiousness impact one's moral self-concept. Finally, using different kinds of self-reported altruism scales, Oda et al. (2014) and Furnham et al. (2016) find that agreeableness, conscientiousness, and openness can explain how altruistically people tend to behave in various situations. Furthermore, self-concept helps people make decisions (Burnett, 1991; Back et al., 2009; Perugini and Leone, 2009). An individual who sees herself as altruistic may be more likely to help someone in need than an individual who does not possess such self-concept. Altruistic self-concept motivates people to maintain and reinforce one's self-image (Bekkers and Wiepking, 2011), put more emphasis on other-regarding behavior (Bénabou and Tirole, 2006) and increases one's warm-glow altruism (Lilley and Slonim, 2014; Tonin and Vlassopoulos, 2014). To be specific, previous studies find that when primed with altruistic self-concept, prosocial individuals are persistently doing good and maintain their positive self-concept (Ariely and Norton, 2008; Stone and Cooper, 2001; Young et al., 2012). These studies attribute the phenomenon to the principle of moral reinforcement, which suggests that individuals will avoid experiencing the disutility of being dissonant with their established self-concept.

To study the effects of personality traits on volunteering decisions, we designed an experiment with 251 participants recruited from Amazon Mechanical Turk (MTurk). In our experiment, participants answered a set of questions and were also given a chance to make

[45]Other authors have used terms such as self-schema or self-image instead of self-concept. Whereas self-schema encompasses how one sees oneself in a particular domain or situation, self-concept is a more comprehensive evaluation of the self that includes the self-schema in all domains.

a promise to volunteer at a later day by answering additional survey questions to benefit a charity. Those who made a promise to volunteer were allowed to keep or renege on their promise. In the questionnaires, we included qualitative and quantitative questions related to altruism which provided us with a measure of altruistic self-concept (the mediator) to build an altruistic self-concept index. The results of our regression analysis show that agreeableness works through altruistic self-concept to influence the probability of making and keeping a promise to volunteer. More specifically, a one standard deviation increase in *Agreeableness* via the index results in an 8.87% increases in the probability of making a promise to volunteer, and renders a 6.30% decrease in the probability of reneging. With respect to openness and conscientiousness, we found weak evidence that they play a role in volunteering decisions. Finally, neither extraversion nor neuroticism affects the decision to make and keep a promise to volunteer.

In our experiment, we did not manipulate variables to "cause" more or less volunteering. Although, in principle, one could manipulate altruistic self-concept perhaps through priming, we did not do that in our experiment. The reason is that no previous studies had considered self-concept as a causal pathway for affecting prosociality. So, there wasn't a good empirical reason to believe that it should be a good candidate for influencing volunteering behavior. In fact, we are the first to consider self-concept as a mediator. Through our analysis in this paper, we are able to identify altruistic self-concept as a causal pathway and show that it mediates the influence of agreeableness on the willingness to make and keep promises to volunteer. This finding is important because it tells us that self-concept is a good candidate for manipulation. By making self-concept silent or salient (e.g., asking subjects to recall past selfless acts), one could affect promises to volunteer.

Recent research in psychology and organizational behavior has shown that personality traits such as agreeableness, openness, and conscientiousness influence prosociality and performance at work.[46] In contrast to previous works, ours is the first paper to link specific

[46]Previous studies have shown that agreeableness and openness predict more prosociality (see Becker et

personality traits *to making and keeping promises* to help. Furthermore, we constructed an index that captures altruistic self-concept and found that it mediates the effects of personality on volunteering decisions. Without the altruistic self-concept index, we did not find any evidence of personality traits affecting the probability of making a promise to volunteer. However, using mediation analysis, we showed that individuals with higher agreeableness are more likely to promise to volunteer, but the effect passes entirely through the mediator. The likelihood of agreeable individuals reneging on a promise to volunteer also passes through altruistic self-concept and is moderated by volunteering experience.

Compared to previous mediators, altruistic self-concept refers to how one describes and evaluates oneself. As mentioned above, one advantage of this mediator over others is that self-concept can be made salient or silent experimentally. For example, one can make altruistic self-concept more salient by asking subjects to recall past selfless acts. Thus, our study not only shows that agreeableness is an important trait for volunteering, but it also informs how this trait works to enhance prosociality. For instance, agreeable individuals who evaluate and describe themselves as altruistic can be more helpful (offer to volunteer) and dependable (keep volunteering promises). Our study results suggest that organizations that rely on volunteers for their services could find ways to strengthen altruistic self-concept, thereby positively influencing prosociality in the workplace.

We begin, in Section 3.2, we describe the experimental study and how we build the altruistic self-concept index. In Section 3.3, we discuss the results. Section 3.4 includes the conclusion and a discussion of the implications of our study.

---

al., 2012; Brown and Taylor, 2015; Kline et al., 2017; Habashi et al., 2016), whereas agreeableness and conscientiousness predict better performance (Barrick and Mount, 2000; Witt et al., 2002; Meyer et al., 2009). In addition, those who score higher on agreeableness and openness tend to contribute more to social causes (Hirsh, 2010; Pavalache-Ilie and Cazan, 2018), while those who score higher on conscientiousness and agreeableness tend to put more effort (Mount and Barrick, 1995) and spend more time at work (Biderman et al., 2008).

3.2 Experimental Design and Procedures

3.2.1 Description of the Study

Our study was conducted online, and it consisted of two surveys scheduled on weekdays three days apart. Each participant was required to answer questions on Day 1 and also participated in a follow-up survey on Day 4. We recruited a total of 284 participants aged 21 to 71 via Amazon Mechanical Turk (MTurk). All participants had an MTurk qualification score of at least 80% and were residents of the United States.[47] 78.09% of the participants were White and 61.75% had some college or higher education.

Each survey took approximately 15 minutes to complete. On Day 1, participants answered questions about exercising, smoking, and drinking habits. They were given a set of Mensa Quizzes (Grosswirth and Salny, 1983), a quantitative time preference survey (Falk et al., 2016), and the Locus of Control scale survey (Rotter, 1966). At the end of these questionnaires, all the participants were directed to the decision-making task, which we explain in the next paragraph. On Day 4, we emailed all participants via MTurk and notified them to complete the follow-up survey. A total of 88.38% of 284 MTurk workers who answered questions on Day 1 participated on Day 4 too.[48] Thus, there were a total of 251 participants (128 males and 123 females) in the sample that we analyzed. The Day 4 follow-up survey included the Big Five Inventory questionnaire (John and Srivastava, 1999), qualitative and quantitative preference-surveys to measure participants' altruism self-concept, and (for a second time) quantitative survey questions to measure their time preference (Falk et al., 2016).

As mentioned above, there was a decision-making task on Day 1. This task consisted of

[47] The vast majority of our participants had obtained at least 90% approval rates. Arechar et al. (2018), and Peer et al. (2014) discuss the importance of recruiting reputable workers. Moreover, to test for reliability of our participants' answers, we administered the time quantitative survey twice on Day 1 and on Day 4. We found that, on average, responses were consistent.

[48] The 33 MTurk workers who did not join our experiment on Day 4 had a lower rate of promise to volunteer compared to those who participated on both dates. However, the difference was not statistically significant. In addition, there were no significant differences with respect to demographic characteristics.

making a promise to answer additional survey questions to benefit a GiveWell[49] charity on Day 4. If on Day 1 participants made a promise to volunteer, in the follow-up survey they were given the option to keep or renege on their promise to volunteer. If they kept their promise, they were directed to a survey which we call the "volunteer survey." The volunteer survey took about 7 minutes to complete.

All participants were informed that they would receive a $2 participation fee for completing the experiment each day. If they finished both experiments, they would be paid a total of $4 upon completion of the follow-up survey on Day 4. Immediately after we reviewed and approved their work, participants received their payment via MTurk. In addition, we made a $2 donation to a GiveWell.org charity for each participant who completed the volunteer survey. The timeline of our experiment is summarized in Table C.3 in the Appendix.

3.2.2 Personality Traits and Altruistic Self-Concept

To measure personality traits, we used a brief version of the Big Five personality scale – the Big Five Inventory (BFI) questionnaire. The BFI consists of 44 short questions (John and Srivastava, 1999) and assesses the five personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. The original Big Five scale (Costa and McCrae, 1992) contains 240 questions; however, the BFI provides measures that are as reliable as the original much longer scale. For altruistic self-concept, we selected 18 qualitative questions from Falk et al.'s (2016) preference survey module to build the altruistic self-concept index.[50] The qualitative questions are divided into three parts. The first part includes participants' self-evaluation of their willingness to share without reciprocation compared to others. The second part includes seven hypothetical questions which ask about

[49] GiveWell.org (http://www.givewell.org/charities/top-charities) is an organization centralizing donations to charities that work to alleviate poverty around the world. According to research, there are three charities that are most in need: Against Malaria Foundation, Schistosomiasis Control Initiative, and Deworm the World Initiative. These are evidence-backed, thoroughly vetted, and underfunded organizations which are shown to be effective programs to alleviate poverty in developing countries.

[50] Falk et al. (2006) have validated the self-reported survey method with behavioral data to ensure that survey responses provide reliable measures.

participants' willingness to share without reciprocation towards a variety of potential recipients. For example, recipients could be *"people in your hometown."* The third part includes ten altruistic tendency questions. They require participants to score the extent to which each question describes them. For example, one question states, *"At work I am only willing to do something for a colleague if I expect that he would do the same for me."* To construct the altruistic self-concept index, we used principal components analysis (PCA) with inputs being standardized scores of the three parts explained above.[51] We define the first component of the PCA as the index, which explains 81.11% of the total variance of these scores. We also added quantitative measures of altruism including the amount of hypothetical money to be sent in a dictator game. We constructed a new score that included the quantitative questions. We later used this combined score to check for the robustness of the mediation mechanism (See Appendix C.1).

3.3 Results

3.3.1 The Effects of Personality Traits on the Promise to Volunteer

To determine whether self-concept is an effective mechanism whereby personality traits influence volunteering decisions, we did a mediation analysis. Figure 3.1 represents the framework for the mediation analysis. The graph shows the effects of one variable onto another. Path c on the graph represents the total effect of the independent variables, X_i on the dependent variable Y (the path does not include the mediator). Path a on the graph (the solid line) represents the effect of the independent variables, X_i, on the mediator, M. Path b on the graph (the dotted line) represents the effect of the mediator, M, on the dependent variable, Y whereas, Path c' (the dashed line) represents the direct effect of the independent variables on Y.

In our statistical models, Path a is derived from regressing the mediator, an index mea-

[51] Standardized scores enabled us to combine the qualitative questions with quantitative questions.

suring altruistic self-concept, on the personality traits, or $M = \beta_0 + aX_i + \epsilon_M$. The effect of M on Y and the effect of X_i on Y are derived from regressing our outcome variables (promise to volunteer or reneging on a promise) on altruistic self-concept and personality traits. That is, $Y = \beta_1 + c'X_i + bM + \epsilon_Y$. The indirect effect of X_i on Y is represented by path ab.

Column (1) in Table 3.1 shows the results of regressing promise to volunteer on personality traits. The estimated coefficients in column (1) represent the log odds of the total effect of traits on promise to volunteer (path c in mediation analysis - see Figure 3.1). None of the estimated coefficients are statistically different from zero.

FIGURE 3.1: MEDIATION ANALYSIS

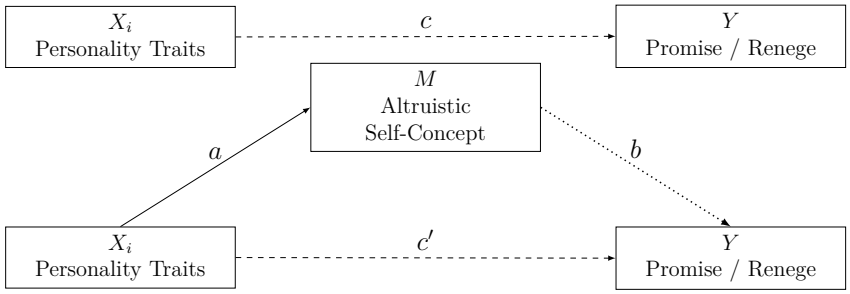


TABLE 3.1: MEDIATION ANALYSIS - PROMISE TO VOLUNTEER

	Path c	Path a	Path b & c'	Path b & c'
	(1)	(2)	(3)	(4)
	Promise	Index	Promise	ME
Index	NA	NA	**0.490*****	[0.105]
Conscientiousness	−0.174	0.167	−0.276	[−0.059]
Neuroticism	0.069	0.148	0.014	[0.003]
Openness	0.174	0.206*	0.098	[0.021]
Extraversion	0.115	−0.090	0.164	[0.035]
Agreeableness	0.111	**0.989*****	−0.359	[−0.077]
N. Obs	251	251	251	
(Pseudo) R-Square	0.013	0.437	0.061	

$*$ p < 0.05; $**$ p < 0.01; $***$ p < 0.001. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Column (2) shows the results of an OLS regression of the index on the Big Five personality traits (path a). *Openness* and *Agreeableness* have a positive effect on the index. Only *Agreeableness* stays significant (p-value < 0.001) after applying the Benjamini-Hochberg

procedure (Benjamini and Hochberg, 1995) for multiple hypotheses testing to control the False Discovery Rate (FDR). Column (3) presents the results of a Logistic regression of promise to volunteer on the index (path b) and the Big Five personality traits (path c'). Only the index has a direct negative effect on promise to volunteer. More specifically, looking at the direct effect only, we can see that increasing the index by one standard deviation results in a 10.5% point increase in the probability of making a promise (see column 4 in Table 3.1 for marginal effects).

The indirect effects of personality traits on the promise to volunteer are summarized in Table 3.1. We did not find a statistically significant indirect effect of *Conscientiousness* via the index. Also, we found that *Openness* does not have statistically significant indirect effect on promise to volunteer via the index (Path ab = 0.051, 95% CI = [-0.005, 0.107], Sobel test p-value = 0.068)[52]. However, *Agreeableness* is inconsistently mediated (path ab = 0.254, 95% CI = [0.119, 0.389], Sobel test p-value = 0.001).[53] Inconsistently mediated means that the signs of coefficient for *Agreeableness* change. The direct effect is negative, but the indirect effect is positive. This means that the mediator is working as a suppressor. In other words, although *Agreeableness* is linked to a lower chance of making a promise, it yields more promises through the altruistic self-concept.

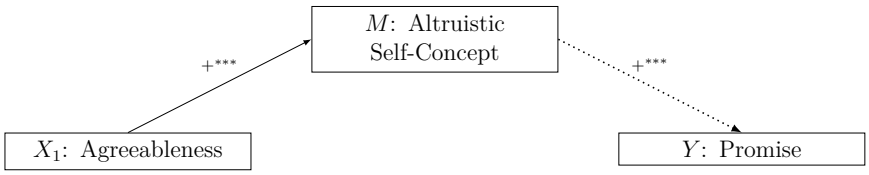
Without the altruistic self-concept index, we do not find any evidence of personality traits affecting the probability of making a promise to volunteer. However, using mediation analysis, we show that individuals with higher *Agreeableness* (cooperative and unselfish) are more likely to promise to volunteer (being more helpful) through the mediator. Specifically, the marginal mediation effect (MME; Barrett et al., 2019) for *Agreeableness* is positive. A one standard deviation increase in *Agreeableness* via the index results in a 10.385% increase

[52]We estimated the confidence interval of our indirect effect using the methodology discussed in Preacher and Hayes (2008).

[53]Inconsistent mediation appears when the sign of path ab is opposite to the sign of path c'. See MacKinnon et al. (2000, 2007) for a list of requirements for inconsistent mediation. Inconsistent mediators can also suggest that there may be confounding variables or variables that may be correlated with two variables of interest. We analyze these issues below.

in the probability of making a promise. These findings suggest that self-concept works as a mechanism whereby personality affects decisions. Figure 3.2 is a graphical summary of the indirect and direct pathways whereby our mediator influences decisions based on the findings in Table 3.1. The solid lines represent the effect of personality traits on the mediator. The dotted line represents the effect of the mediator on promise to volunteer controlling for personality traits.

FIGURE 3.2: PROMISE TO VOLUNTEER



Note: $*$ p < 0.05; $**$ p < 0.01; $***$ p < 0.001. Figure 2 is a graphical summary of column (2) and (3) in Table 3.1. In this figure we find *Agreeableness* is the only personality traits has indirect effect on the probability of making a promise to volunteer via the index.

Previous experimental studies have documented differences in prosocial behaviors between men and women. Andreoni and Vesterlund (2001) and Cox and Deck (2006) showed that men are more responsive to the cots of altruism and women are more altruistic toward their closer networks, respectively. With respect to altruistic self-concept, in a meta-analysis on how sex differences relate to altruism, Eagly and Crowley (1986) found that chivalry and heroism drive helping behavior. In our study, sex may influence both our dependent variables and the mediator. This means that sex can be a confounder in our model and thus, controlling for sex may be important. In addition, sex can interact with personality to influence behaviors. For example, more agreeable males could have a stronger altruistic self-concept than females, while more agreeable males who have higher altruistic self-concept may be more likely to volunteer than their female counterparts. Thus, sex could be a moderator for the paths $X \to M$ or $M \to Y$.[54] Figure 3.3 shows the possible effects of sex as a confounder

[54]A moderator is a variable that could cause the relationship of the independent variable(s) on a dependent variable to be different across different values of the moderator. See Hayes (2015) for an explanation of

C and as a moderator W_1 and W_2. We extend our mediation model to include a sex dummy (*Female*) as moderator and as confounder. Our regression results show that neither *Female* nor any of the interaction terms are statistically different from zero (see columns (1) - (4) in Table C.4in the Appendix). The latter means that *Female* does not moderate the effect of personality on promise to volunteer.

Volunteering experience can also influence both altruistic self-concept (Bekkers and Wiepking, 2011; Piliavin and Callero, 1991) and the willingness to volunteer again (Clary and Snyder, 1999). Experience can also interact with personality to affect prosocial behaviors (King et al., 2014). Thus, in a manner similar to sex, volunteering experience could be a confounder and/or a moderator. We added to our regression a dummy representing volunteering experience (*VolExp*) and interacted this variable with the Big Five personality traits (see Table C.4 columns (5) - (8)). The coefficient for *VolExp* representing the direct effect of volunteering experience on the likelihood of making a promise to volunteer is positive and significantly different from zero (coeff = 1.113, p-value < 0.001, see column (5)). However, none of the interaction terms of this variable with the Big Five are different from zero.

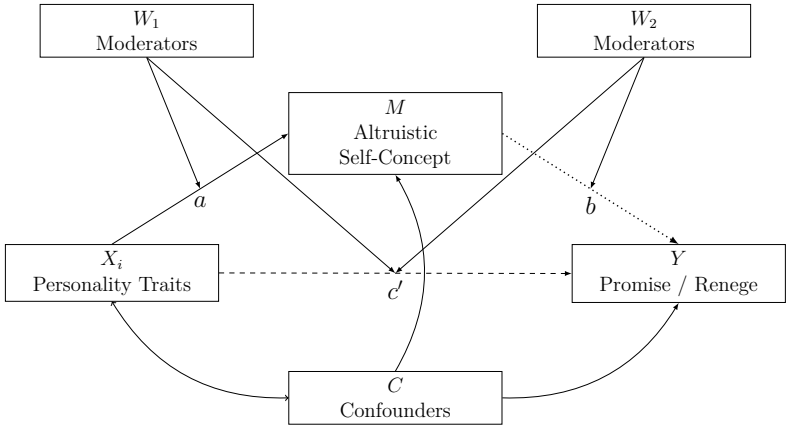
Finally, regressing the likelihood of making a promise to volunteer on the interaction terms *Female*×*Index* and *VolExp*×*Index* shows that these variables don't affect outcomes as W_2 moderators - see Figure 3.3 for W_2 and Table C.6 (columns 1 - 4) for empirical results.

The mediation analysis including *VolExp* yields almost identical results to those shown in Table 3.1. That is, *Agreeableness* is inconsistently mediated (path $a_1b = 0.233$, 95% CI = [0.092, 0.375], Sobel test p-value = 0.002).[55] The MME for *Agreeableness* is still positive. Specifically, one standard deviation increase in *Agreeableness* via the index results in a 8.869% increases in the probability of making a promise. Figure 3.4 graphically summarizes our findings. The dashed line represents the (direct) effect of confounder on promise to volunteer

moderated mediation and Hayes (2018) for an extension of the moderated mediation literature to more than one moderator. A more detailed explanation of how we integrate moderators and confounders in our model can be found in Appendix C.2.

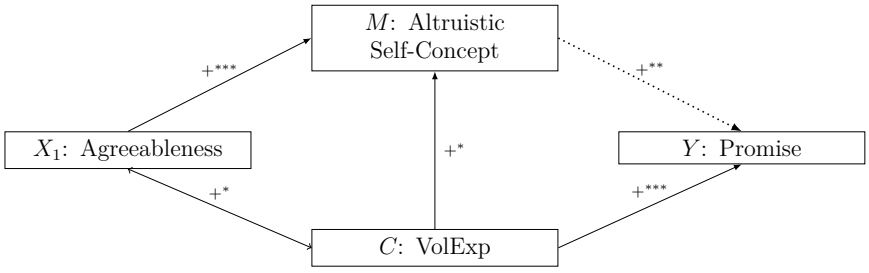
[55] The findings here based on results in columns (5) and (6) in Table C.4. See Appendix C.2 for an explanation of path a_1b in moderated mediation.

FIGURE 3.3: MEDIATION ANALYSIS WITH MODERATORS



controlling for the mediator. Figure 3.5 summarizes the MME of all personality traits on the probability of making a promise to volunteer. As previously noted, *Agreeableness* is the only personality trait that has statistically significant indirect effect.

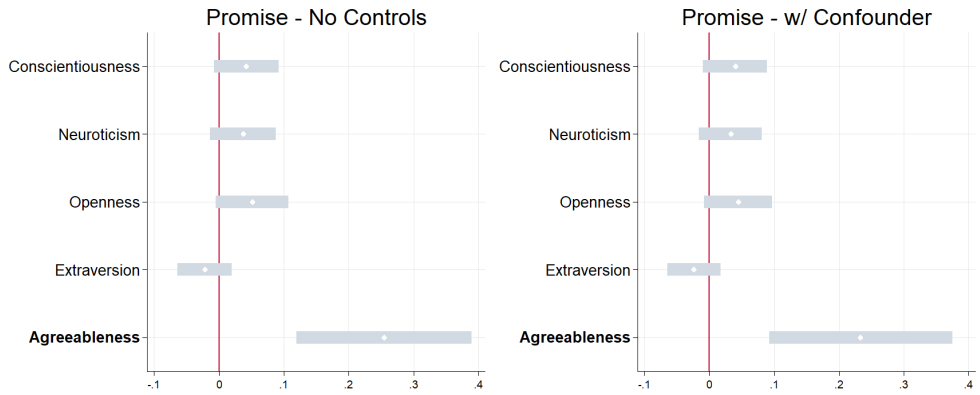
FIGURE 3.4: MODERATED MEDIATION ON PROMISE TO VOLUNTEER



Note: $*$ p < 0.05; $**$ p < 0.01; $***$ p < 0.001. Figure 4 is a graphical summary of column (5) and (6) in Table C.4. We find *Agreeableness* contributes to the probability of making a promise to volunteer indirectly via the index, while *VolExp* has positive direct effect on the dependent variable.

Although not statistically significant after applying the Benjamini-Hochberg procedure (p-value = 0.084), the finding that agreeableness has negative **direct** impact on promising to volunteer is a little surprising to us. Indeed, previous studies (Becker et al., 2012; Ben-Ner et al., 2008; Volk et al., 2011, 2012; etc.) have found that agreeableness is often positively correlated with altruistic behaviors. The discrepancy might be due to the fact that we asked people to make a promise to volunteer in the future. To see if this would explain the discrepancy, we tested the effect of personality traits on effort donations (not a promise). A

FIGURE 3.5: MARGINAL MEDIATION EFFECTS - PROMISE TO VOLUNTEER



Note: The white dots on the figure depict the marginal mediation effects of personality traits on the probability of making a promise to volunteer. The shaded blocks represent 95% confidence intervals.

group of 134 MTurk workers participated in an experiment identical to this one, but they were asked the following question *"Would you be willing to answer additional questions to benefit a charity* ***now****?"*. We ran regressions similar to those in Table 3.1 and columns (5) and (6) in Table C.4, where the dependent variable was "volunteer now" rather than "make a promise to volunteer later." We found that *Agreeableness* had a strong positive direct effect on effort donations (See Appendix C.2). This suggests that the discrepancy with other studies is due to asking for a promise. Making a promise to do something for others in the future adds an element of uncertainty and thus, agreeable people would rather not promise to volunteer if they were not sure that they could help on a future date.[56]

3.3.2 The Effects of Personality Traits on Reneging

In Section 3.3.1, we found that altruistic self-concept mediates the effect of some personality traits on the promise to volunteer. In this section, we test whether the index could also mediate the effects of personality traits on the probability of reneging on a promise to

[56] As recommended by a referee, we reconsidered whether the study had enough power. We conducted a post-hoc power analysis utilizing the methodology in Schoemann et al. (2017). The results showed that our sample size was large enough to support the main findings of our paper.

volunteer. Similar to the previous analysis, we construct an OLS regression with the index as the dependent variable and the Big Five personality traits as independent variables. We also run a logit regression with the probability of reneging as the dependent variable. As previously done, we add *Female* and *VolExp* as potential confounds and moderators (See Figure 3.3 and Table C.5 in the Appendix).

We summarize some regression results in Table 3.2.[57] We find that *VolExp* interacts with *Agreeableness* to influence the index (see column (1)). The coefficient of this interaction term is negative (OLS coef= -0.873, p-value = 0.001), which is opposite to the effect of the variable *Agreeableness* itself on the index.[58] After using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) for multiple hypotheses testing to control the False Discovery Rate (FDR), the p-values of *Agreeableness*, and its interaction with *VolExp* remain significant (p-value < 0.05). The moderated mediation analysis shows that *Agreeableness* is inconsistently mediated (path a_1b = -0.640, 95% CI = [-1.121, -0.160], Sobel test p-value = 0.003). In other words, the mediator suppresses the effect of this personality trait on reneging. Likewise, the interaction of *Agreeableness* and volunteering experience is also inconsistently mediated (path a_3b = 0.316, 95% CI = [0.039, 0.592], Sobel test p-value = 0.022).[59] Despite the interaction term having a positive indirect effect on reneging, the overall indirect effects of *Agreeableness* suggest that more agreeableness leads to lower reneging through the index (being more dependable).

When the variable *VolExp* is included in the regressions as moderator and confounder, *Conscientiousness* has a direct negative effect on the probability of reneging (see column (8)). Intuitively, conscientiousness stands for "responsibility," which demotivates people to renege on a promise. However, the effect is **not** significant after correcting for multiple comparisons.

[57] The table includes two columns transposed from Table C.5 in the Appendix.

[58] We test if *Female* moderates the effect of personality traits on the index via its interaction with *Extraversion* (see columns (3) and (9) in Table C.5). However, the indirect effect of this interaction term is not different from zero (column 3: 95% CI = [-0.069, 0.403]; column 9: 95% CI = [-0.736, 0.073]). Therefore, we reject *Female* is an effective moderator or confounder in explaining the probability of reneging on a promise.

[59] See Appendix C.2 for Path a_1b and Path a_3b in moderated mediation. The indirect effect of *Agreeableness* on reneging via the moderated mediator is positive and significant: $a_1b + a_3bW = -0.640 + 0.316 \times$ *VolExp*.

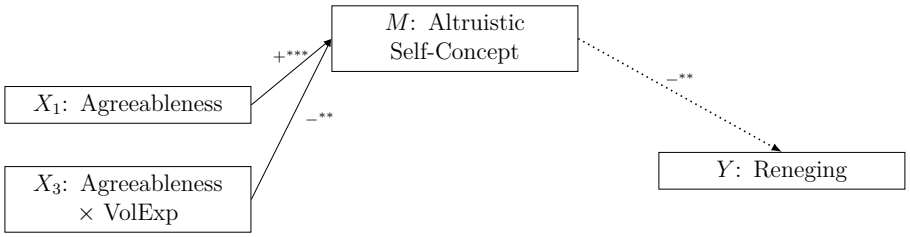
Table 3.2: Moderated Mediation Analysis - Renege on a Promise

	(7) Index	(8) Renege
Index	NA	**-0.898****
Conscientiousness	−0.089	−1.150
Neuroticism	0.555*	−0.338
Openness	0.043	0.710
Extraversion	0.051	−0.071
Agreeableness	**1.567*****	0.628
VolExp	−0.179	−0.988
Conscientiousness × VolExp	0.586*	0.982
Neuroticism × VolExp	−0.273	−0.242
Openness × VolExp	0.520	0.302
Extraversion × VolExp	−0.049	0.415
Agreeableness × VolExp	**-0.873****	−0.742
N. Obs	92	92
(Pseudo) R-Square	0.572	0.218
AIC	371.730	

$*$ p $<$ 0.05; $**$ p $<$ 0.01; $***$ p $<$ 0.001. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Figure 3.6 displays the graphical pathways (see also Table 3.2 for the estimated coefficients).[60] In Figure 3.7, we depict the moderated MME of all personality traits on the probability of reneging.

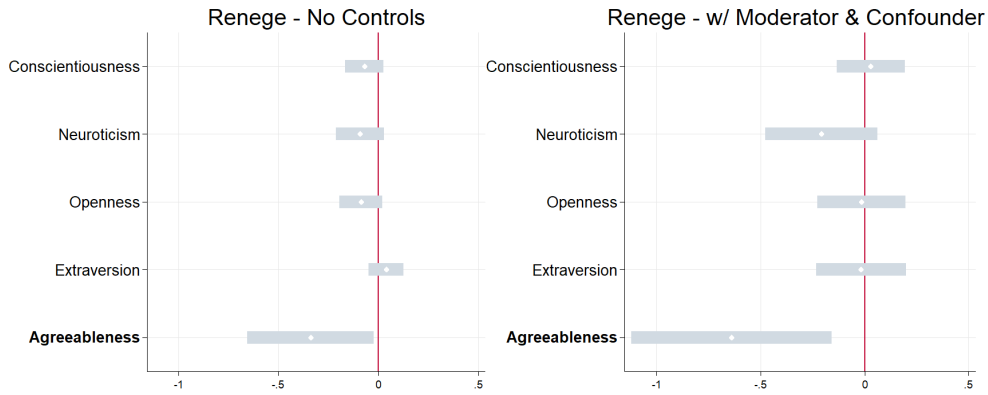
Figure 3.6: Moderated Mediation in Reneging on a Promise



Note: $*$ p $<$ 0.05; $**$ p $<$ 0.01; $***$ p $<$ 0.001. Figure 6 is a graphical summary of Table 3.2. We find both *Agreeableness* and its interaction with *VolExp* contribute to the probability of reneging indirectly via the index.

[60] *Female* and *VolExp* don't affect outcomes as W_2 moderators (see Figure 3.3 and Table C.6 (columns 5 - 8) in the Appendix).

Figure 3.7: Marginal Mediation Effects - Reneging on a Promise



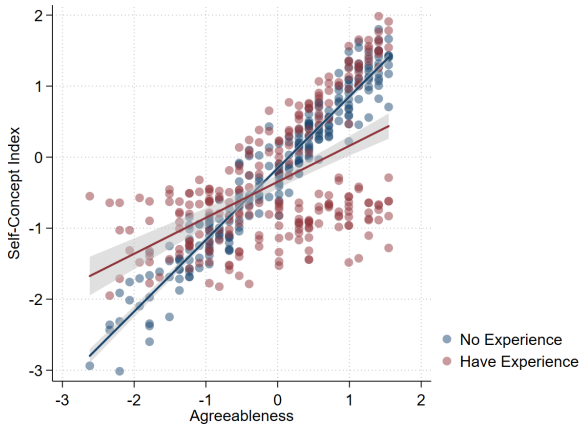
Note: The white dots on the figure depict the marginal mediation effects of personality traits on the probability of reneging on a promise to volunteer. The shaded blocks represent 95% confidence intervals.

All in all, with respect to reneging on a promise to volunteer, we find that the effect of personality traits on the index depends on whether subjects have volunteering experience. Figure 3.8 is a plot showing the relationship between *Agreeableness* and the index by experience group (have experience and no experience). The steeper slope for the inexperienced group compared to the experienced group reflects a stronger influence of *Agreeableness* on altruistic self-concept in the former group. Intuitively, individuals with higher agreeableness, regardless of whether they have volunteering experience or not, would be more likely to have higher altruistic self-concept (positive slope for both groups) and therefore keep their promise. However, for those who have no volunteering experience, the effect of altruistic self-concept is stronger the more agreeable they are.

3.4 Conclusion and Discussion

In this paper, we study the effects of personality traits on the willingness to make and keep a promise to volunteer. We designed an online experiment with Amazon Mechanical Turk workers who were tasked to fill out two surveys three days apart. On the first day (Day 1), participants answered a set of questions and were given an opportunity to make a promise to volunteer three days later (on Day 4). Volunteering consisted of answering additional

Figure 3.8: Agreeableness × Volunteering Experience



survey questions to benefit a charity. On Day 4, all participants answered the Big Five personality questionnaire and those who had promised to volunteer were given the option to keep or renege on their promises. We also collected self-evaluations and quantitative measures of altruism that we used to build an altruistic self-concept index. We utilized this index as a mediator to get a better understanding of the mechanisms whereby personality traits influence volunteering decisions.

Our regression analysis shows that agreeableness positively and indirectly influences one's likelihood of making a promise to volunteer through the altruistic self-concept index. More specifically, we find that more agreeable individuals are more likely to make a promise to help via the index. With respect to keeping promises, we find that altruistic self-concept mediates agreeableness, and volunteering experience moderates the effect. In other words, more agreeable individuals are less likely to renege on their promise to volunteer. However, compared to inexperienced individuals, those with volunteering experience score lower in altruistic self-concept, the more agreeable they are. Thus, altruistic self-concept has a stronger mediating effect when agreeable individuals are inexperienced than when they are experienced.

Whereas other researchers have studied the relationship between personality and prosocial behaviors (see Jiang and Capra, 2018; Capra et al., 2013; Borghans et al., 2009; Almlund et al., 2011; Heckman et al., 2019; Rustichini et al., 2016; etc.), ours is the first paper that

links personality traits to *making and keeping promises to help*. Furthermore, our mediation analysis helped us understand the mechanisms whereby personality influences decisions. We find that altruistic self-concept mediates the effects of personality on volunteering. Without this index as a mediator, we did not find any evidence of personality traits affecting the probability of making a promise to volunteer. However, using mediation analysis, we showed that individuals with higher *Agreeableness* are more likely to promise to volunteer, but the effect passes through the altruistic self-concept. The likelihood of agreeable individuals reneging on a promise to volunteer also passes through the index. These findings have important implications. Although it is not possible to change one's personality, we could influence altruistic self-concept. For example, we could ask agreeable subjects to recall situations in which they were generous to others. By affecting altruistic self-concept through priming we could increase individuals' willingness to make promises to volunteer. According to our findings, the effect on keeping promises would be larger among those with little or no experience volunteering[61].

Our study has broader implications for the non-profit sector. The willingness to make a promise to volunteer can be thought of as helpfulness, and the willingness to keep a promise as dependability. Helpfulness and dependability are two critical traits often linked to good performance of volunteers. In the US, for example, hospitals and hospices would not be able to offer high quality and personalized health care without raising costs, if they did not have enough help from reliable volunteers (Hotchkiss et al., 2008, 2014; Kentner et al., 2003; Wells et al., 1990). Volunteers not only help the health care system financially, but they also "add a human touch to the technical aspect of care by contributing to the happiness and comfort of patients" (Hotchkiss et al., 2008, page 2). Thus, from the perspective of administrators and patient advocates, identifying the right volunteer workforce matters. Our study suggests

[61] In our experiment, we did not manipulate variables to "cause" more or less volunteering. In fact, mediation analysis is used precisely for the purpose of identifying causal pathways as noted in Imbens (2020), "The value of the mediation analysis is that it sheds light on the causal pathways. The challenge is that it requires us to identify a number of constituent causal effects" (p. 1146). Through mediation analysis, we were able to identify altruistic self-concept as a causal pathway for influencing prosociality.

that agreeable individuals who evaluate and describe themselves as altruistic can be more helpful and dependable. Organizations that rely on volunteers for their services could find ways to strengthen altruistic self-concept, thereby positively influencing prosociality in the workplace.

Appendices

A Chapter 1 Appendix

A.1 Mathematical Proofs

Proofs of Proposition I to III With Function Form:

In this section we use the same function form as in Section 1.2.2 (ability is known). We consider When there is only 2 values of utility, set g such that $\theta_L q < g < \theta_H q$, the optimal q under each is:

$$a\theta_H + \frac{1}{2}\theta_H \mu(\theta_H q_H - g)^{-\frac{1}{2}} = q_H$$
$$a\theta_L + \frac{1}{2}\theta_L \mu\lambda(g - \theta_L q_L)^{-\frac{1}{2}} = q_L$$

SOC:

$$-1 - \frac{1}{4}\theta_H \mu(\theta_H q_H - g)^{-\frac{3}{2}} < 0$$
$$-1 + \frac{1}{4}\theta_L \mu\lambda(\theta_L q_L - g)^{-\frac{3}{2}} < 0$$

then the expected utility $p = \text{prob}(\theta = \theta_H)$

$$\underset{g\geq 0}{max} \quad p[a\theta_H q_H + bg + \mu(\theta q - g)^{\frac{1}{2}} - \frac{{q_H}^2}{2}] + (1-p)[a\theta_L q_L - \mu\lambda(g - \theta_L q)^{\frac{1}{2}} - \frac{{q_L}^2}{2}]$$

FOC: [g]

$$0 = pb - p\mu(\theta_H q_H - g)^{-\frac{1}{2}}\frac{1}{2} - (1-p)\mu\lambda(g - \theta_L q_L)^{-\frac{1}{2}}\frac{1}{2}$$
$$+ p[a\theta_H - q_H + \mu(\theta_H q - g)^{-\frac{1}{2}}\frac{\theta_H}{2}]\frac{dq_H}{dg} + (1-p)[a\theta_L - q_L - \mu\lambda\theta_L(g - \theta_L q)]^{-\frac{1}{2}}\frac{1}{2}]\frac{dq_L}{dg}$$

SOC:

$$\frac{p\mu}{4}(\theta_H q_H - g)^{-\frac{3}{2}}(\theta_H \frac{dq_H}{dg} - 1) + \frac{(1-p)\mu\lambda}{4}(1 - \theta_L \frac{dq_L}{dg}^{-\frac{1}{2}}$$
$$= -\frac{p}{\theta_H}\frac{dq_H}{dg} + \frac{1-p}{\theta_L}\frac{dq_L}{dg} < 0$$

$\Rightarrow \frac{1-p}{\theta_L} < \frac{p}{\theta_H} \Rightarrow \frac{\theta_H}{\theta_L} < \frac{p}{1-p} \Rightarrow \theta_H < \frac{p}{1-p}\theta_L$ is held (see Wu et al., 2008 for myopic optimization).

Then, the g, q_H, q_L is determined by:

$$\begin{cases} 2pb = p\mu(\theta_H q_H - g)^{-\frac{1}{2}} + (1-p)\mu\lambda(g - \theta_L q_L)^{-\frac{1}{2}} \\ \frac{q_H}{\theta_H} = a + \frac{\mu}{2}(\theta_H q_H - g)^{-\frac{1}{2}} \\ \frac{q_L}{\theta_L} = a + \frac{\mu\lambda}{2}(g - \theta_L q_L)^{-\frac{1}{2}} \end{cases}$$

1) b :

$$2p = -\frac{1}{2}p\mu(\theta_H q_H - g)^{-\frac{3}{2}}(\theta\frac{dq_H}{dg} - 1)\frac{dg}{db} - \frac{1}{2}(1-p)\mu\lambda(g - \theta_L q_L)^{-\frac{3}{2}}(1 - \theta_L\frac{dq_L}{dg})\frac{dg}{db}$$

$$\Rightarrow \quad 2p = -SOC \times \frac{dg}{db}$$

as: $SOC < 0$

$$\Rightarrow \quad \frac{dg}{db} > 0$$

2) λ:

$$0 = -\frac{1}{2}p\mu(\theta_H q_H - g)^{-\frac{3}{2}}(\theta\frac{dq_H}{dg} - 1)\frac{dg}{d\lambda} - \frac{1}{2}(1-p)\mu\lambda(g - \theta_L q_L)^{-\frac{3}{2}}(-\theta_L\frac{dq_L}{dg}\frac{dg}{d\lambda} - \theta_L\frac{dq_L}{d\lambda} + \frac{g}{\lambda})$$

$$\Rightarrow \quad 0 = SOC \times \frac{dg}{d\lambda} + \frac{(1-p)\mu\lambda}{2}(g - \theta_L q_L)^{-\frac{3}{2}}\theta_L\frac{dq_L}{d\lambda}$$

by FOC: $\frac{1}{\theta_L}\frac{dq_L}{d\lambda} = \frac{\mu}{2}(g - \theta_L q_L)^{-\frac{1}{2}} + \frac{\mu\lambda}{4}(g - \theta_L q_L)^{-\frac{3}{2}}\frac{q_L}{\lambda}\theta_L$

then, $[\frac{1}{\theta_L} - \theta_L\frac{\mu\lambda}{4}(g - \theta_L q_L)^{-\frac{3}{2}}]^{-1}\frac{\mu}{2}(g - \theta_L q_L)^{-\frac{3}{2}} = \frac{dq_L}{d\lambda}$

as: $\frac{1}{\theta_L} - \theta_L\frac{\mu\lambda}{4}(g - \theta_L q_L)^{-\frac{3}{2}} > 0; \quad \frac{dq_L}{d\lambda}; \quad SOC < 0$

$$\Rightarrow \quad \frac{dg}{d\lambda} < 0$$

3) μ :

$$0 = SOC \times \frac{dq}{d\lambda} - \frac{1}{2}p\mu(\theta_H q_H - g)^{-\frac{3}{2}}\theta_H \frac{dq_H}{d\mu} + \frac{1}{2}(1-p)\mu\lambda(g - \theta_L q_L)^{-\frac{3}{2}}\theta_L \frac{dq_L}{d\mu}$$
$$+ p(\theta_H q_H - g)^{-\frac{1}{2}} + (1-p)\lambda(g - \theta_L q_L)^{-\frac{1}{2}}$$

by FOC: $\frac{dq_H}{d\mu}\frac{1}{\theta_H} = \frac{1}{2}(\theta_H q_H - g)^{-\frac{1}{2}} - \frac{\mu}{4}(\theta_H q_H - g)^{-\frac{3}{2}}\theta_H \frac{dq_H}{d\mu} \Rightarrow \frac{dq_H}{d\mu} > 0$

$\frac{dq_L}{d\mu}\frac{1}{\theta_L} = \frac{\lambda}{2}(-\theta_L q_L + g)^{-\frac{1}{2}} + \frac{\mu\lambda}{4}(-\theta_L q_L + g)^{-\frac{3}{2}}\theta_L \frac{dq_L}{d\mu} \Rightarrow \frac{dq_L}{d\mu} > 0$

$$\Rightarrow \quad 0 = -SOC \times \frac{dg}{d\mu} + p[(\theta_H q_H - g)^{-\frac{1}{2}} - \frac{\mu}{2}(\theta_H q_H - g)^{-\frac{3}{2}}\theta_H \frac{dq_H}{d\mu}]$$
$$+ (1-p)[\lambda(g - \theta_L q_L)^{-\frac{1}{2}} + \frac{\mu\lambda}{2}(g - \theta_L q_L)^{-\frac{3}{2}}\theta_L \frac{dq_L}{d\mu}]$$
$$= -SOC \times \frac{dq}{d\mu} + \frac{dq_H}{d\mu}\frac{1}{\theta_H} + \frac{dq_L}{d\mu}\frac{1}{\theta_L}$$
$$\Rightarrow \quad -\frac{dq_H}{d\mu}\frac{1}{\theta_H} - \frac{dq_L}{d\mu}\frac{1}{\theta_L} = -SOC \times \frac{dg}{d\mu}$$
$$\Rightarrow \quad \frac{dg}{d\mu} < 0$$

A.2 Figures

FIGURE A.1: ACCURACY RATE COMPARISONS BETWEEN GROUPS

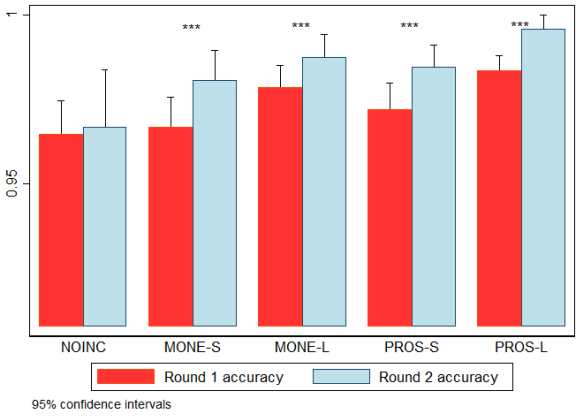
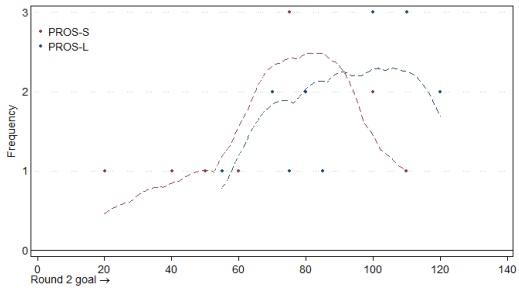


FIGURE A.2: FREQUENCY OF ROUND 2 GOAL BETWEEN PROS-S AND PROS-L

(a)



(b)

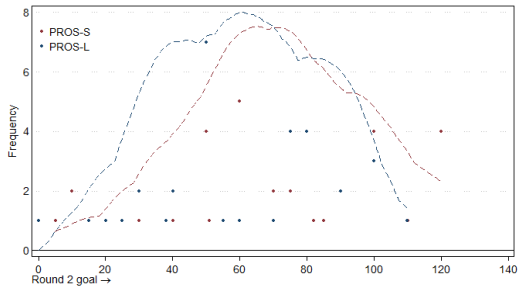
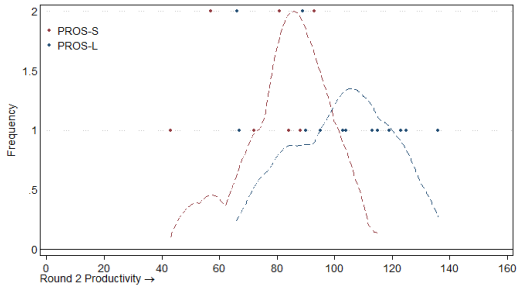


FIGURE A.3: FREQUENCY OF ROUND 2 OUTPUT BETWEEN PROS-S AND PROS-L

(a)



(b)

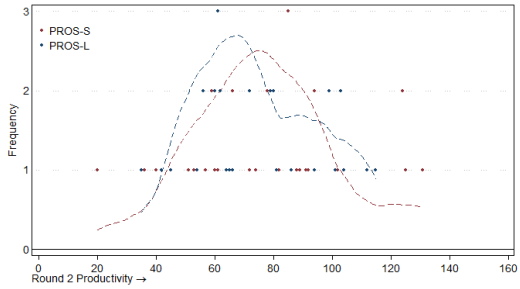


FIGURE A.4: ROUND 2 GOAL-ACHIEVEMENT RATE BETWEEN PROS-S AND PROS-L

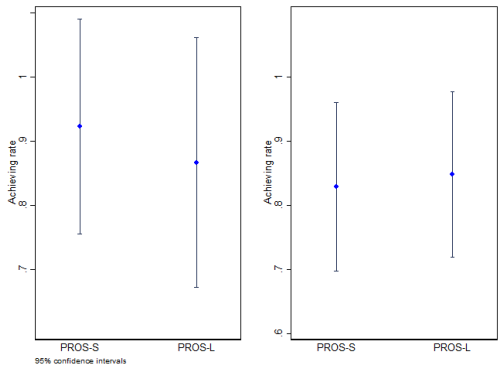


FIGURE A.5: EARNINGS BY GROUP

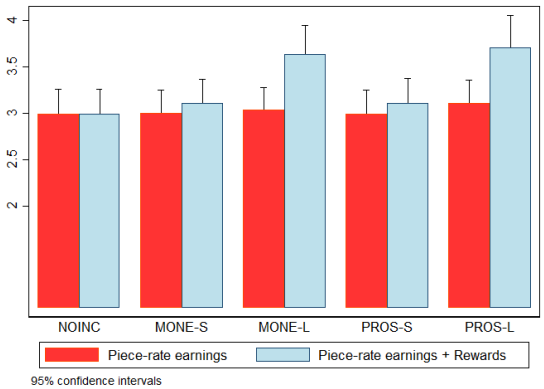


FIGURE A.6: ROUND 2 GOAL AND TASKS COMPLETED

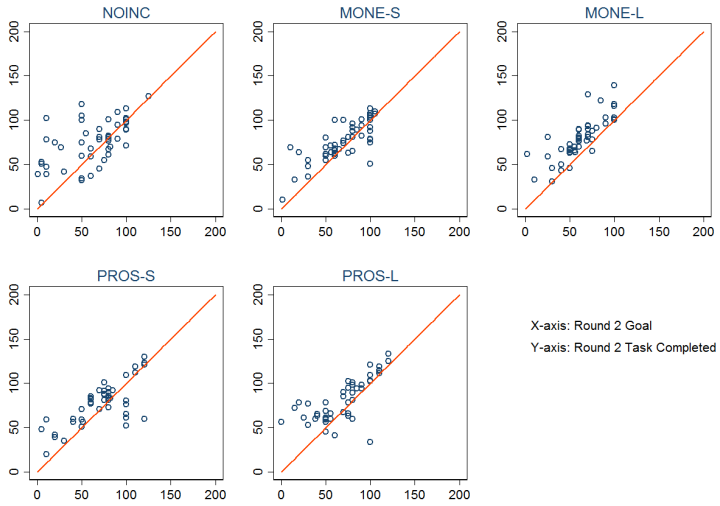
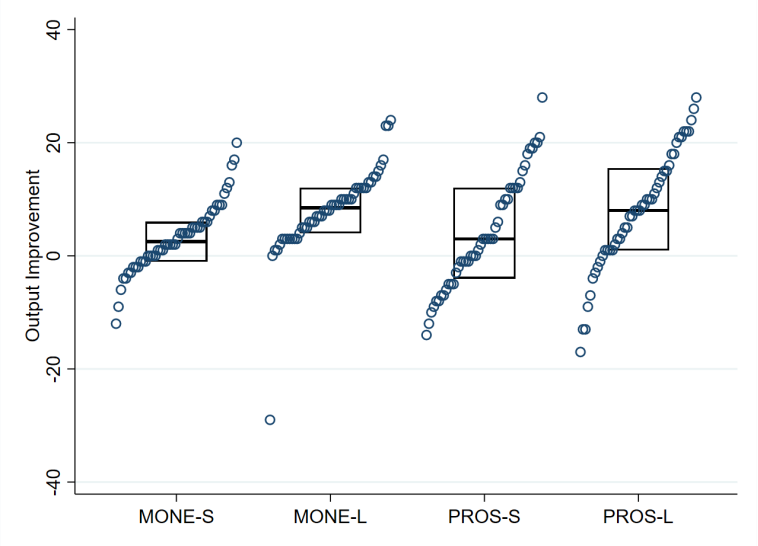
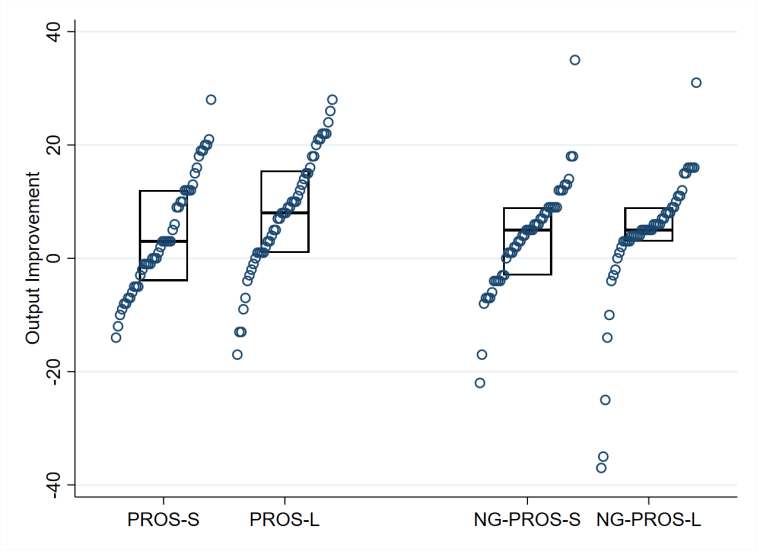


FIGURE A.7: OUTPUT IMPROVEMENT DISTRIBUTION COMPARISONS



(a) monetary vs. prosocial



(b) prosocial with goal vs. prosocial without goal

A.3 Tables

Table A.1: Treatment effects between groups (total output)

	NOINC vs. MONE-S		NOINC vs. MONE-L		NOINC vs. PROS-S		NOINC vs. PROS-L	
	(1) Complete'	(2) Complete'	(3) Complete'	(4) Complete'	(5) Complete'	(6) Complete'	(7) Complete'	(8) Complete'
Treatment	−0.894	−0.451	−3.574	−1.375	−2.021	0.551	−1.875	2.396
	(4.443)	(4.764)	(4.377)	(5.581)	(4.227)	(5.030)	(4.164)	(4.407)
Round2	−0.042	−0.042	−0.042	−0.042	−0.042	−0.042	−0.042	−0.042
	(1.644)	(1.644)	(1.644)	(1.722)	(1.612)	(1.665)	(1.612)	(1.612)
Treatment × Round2	2.362	2.362	7.682***	7.552***	3.354*	3.932**	7.250***	7.250***
	(1.887)	(1.887)	(2.981)	(2.077)	(2.137)	(2.201)	(2.258)	(2.258)
Controls	N	Y	N	Y	N	Y	N	Y
N. Obs	196	196	196	194	192	188	192	192
R-Square	0.001	0.045	0.014	0.040	0.003	0.064	0.014	0.149

Notes: $*$ p $<$ 0.1; $**$ p $<$ 0.05; $***$ p $<$ 0.01. Dependent variable is the number of task completed (total output). Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between NOINC (control group) and MONE-S (small reward with monetary incentive). Columns (3) and (4) are comparisons between NOINC and MONE-L (large reward with monetary incentive). Columns (5) and (6) are comparisons between NOINC and PROS-S (small reward with prosocial incentive). Columns (7) and (8) are comparisons between NOINC and PROS-L (large reward with prosocial incentive). Standard errors cluster at the individual level.

Table A.2: Treatment effects comparisons by reward sizes (total output)

	MONE-S vs. MONE-L		PROS-S vs. PROS-L	
	(1)	(2)	(3)	(4)
	Complete'	Complete'	Complete'	Complete'
MONE-L	−2.680	−1.286		
	(4.477)	(4.831)		
PROS-L			0.146	−2.313
			(4.225)	(5.266)
Round2	2.230***	2.230***	3.313**	3.891***
	(0.885)	(0.887)	(1.389)	(1.431)
MONE-L × Round2	5.320***	5.190***		
	(1.456)	(1.517)		
PROS-L × Round2			3.896**	3.317*
			(2.179)	(2.219)
Controls	N	Y	N	Y
N. Obs	200	198	192	188
R-Square	0.016	0.064	0.018	0.113

Notes: $*$ p < 0.1; $**$ p < 0.05; $***$ p < 0.01. Dependent variable is the number of task completed (total output). Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between MONE-S (small reward monetary incentive) and MONE-L (large reward monetary incentive). Columns (3) and (4) are comparisons between PROS-S (small reward prosocial incentive) and PROS-L (large reward prosocial incentive). Standard errors cluster at the individual level.

Table A.3: Treatment effects comparisons by incentive types (total output)

	MONE-S vs. PROS-S		MONE-L vs. PROS-L	
	(1)	(2)	(3)	(4)
	Complete'	Complete'	Complete'	Complete'
PROS-S	−1.127	0.190		
	(4.172)	(5.308)		
PROS-L			1.699	−0.211
			(4.126)	(4.489)
Round2	2.320***	2.320**	7.640***	7.510***
	(0.872)	(0.914)	(1.152)	(1.175)
PROS-S × Round2	0.993	1.571		
	(1.745)	(1.711)		
PROS-L × Round2			−0.432	−0.302
			(1.966)	(1.997)
Controls	N	Y	N	Y
N. Obs	196	192	196	194
R-Square	0.004	0.055	0.029	0.089

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Dependent variable is the number of task completed (total output). Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Columns (1) and (2) are comparisons between MONE-S (small reward monetary incentive) and PROS-S (small reward prosocial incentive). Columns (3) and (4) are comparisons between MONE-L (large reward monetary incentive) and PROS-L (large reward prosocial incentive). Standard errors cluster at the individual level.

Table A.4: Goal-Setting's effect on Performance Improvement (total output and total goal-achievement dummy)

	NOINC	MONE-S	MONE-L	PROS-S	PROS-L
	(1)	(2)	(3)	(4)	(5)
	ΔComplete'	ΔComplete'	ΔComplete'	ΔComplete'	ΔComplete'
Round2 Goal	0.010	−0.030	0.304**	0.031	−0.051
	(0.087)	(0.045)	(0.139)	(0.085)	(0.094)
Round2 Goal × Achieve'	0.110**	0.020	−0.229**	0.108*	0.215***
	(0.067)	(0.037)	(0.110)	(0.080)	(0.082)
Controls	Y	Y	Y	Y	Y
N. Obs	47	50	49	46	48
R-Square	0.169	0.246	0.426	0.300	0.433

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Control variables include concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. The table display how Round 2 goals and its interaction with the total goal-achievement affect participants' total output improvement in each group. Standard errors cluster at the individual level.

Table A.5: Goal-setting's effect on performance improvement with mission matched or mismatched (total output)

	PROS-S	PROS-L	PROS-S & PROS-L
	(1)	(2)	(3)
	ΔComplete'	ΔComplete'	ΔComplete'
Round2 Goal	0.029	−0.108	−0.003
	(0.085)	(0.106)	(0.041)
Round2 Goal × Achieve'	0.106*	0.203**	0.110***
	(0.081)	(0.087)	(0.037)
Matched	0.801	7.012*	0.868
	(3.749)	(4.334)	(3.209)
PROS-L			0.252
			(2.464)
PROS-L × Matched			5.734*
			(4.397)
Controls	Y	Y	Y
N. Obs	46	48	94
R-Square	0.301	0.486	0.351

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Standard errors cluster at the individual level. This table displays how being mission matched and its interaction with the large reward size affect participants' total output improvement. Standard errors cluster at the individual level.

Table A.6: Mediation Analysis of the Mechanism (total output)

	(1)	(2)
	Round2 Goal	ΔComplete'
Round2 Goal	NA	0.105***
		(0.042)
Matched	10.114	1.989
	(8.925)	(3.699)
PROS-L	−4.436	2.885
	(9.251)	(2.700)
PROS-L × Matched	23.435**	4.330
	(12.950)	(4.684)
Controls	Y	Y
N. Obs	80	80
R-Square	0.265	0.324

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. All columns use OLS regression. This table displays the effect of large rewards interacts with being mission matched on one's total output improvement via the self-chosen goal. It only includes participants who assumed that they achieve their goals. Standard errors cluster at the individual level.

Table A.7: Treatment effects between productivity levels

Panel A	NOINC vs. MONE-S		NOINC vs. MONE-L		NOINC vs. PROS-S		NOINC vs. PROS-L	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Complete	Complete	Complete	Complete	Complete	Complete	Complete	Complete
Treatment	−0.309	0.750	−0.444	4.608	−0.959	−1.348	2.868	−2.579
	(3.450)	(4.067)	(3.304)	(5.847)	(4.292)	(4.751)	(3.409)	(4.066)
Round2	−1.458	3.000***	−1.458	3.000***	−1.458	3.000***	−1.458	3.000***
	(2.941)	(1.161)	(2.783)	(1.154)	(2.900)	(1.226)	(2.896)	(1.148)
Treatment × Round2	4.898*	0.080	9.569***	5.000***	6.719**	1.304	7.197**	6.960***
	(3.159)	(1.657)	(3.269)	(1.627)	(3.583)	(2.417)	(3.605)	(2.178)
High-productivity	N	Y	N	Y	N	Y	N	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
N. Obs	98	98	102	92	94	94	94	98
R-Square	0.206	0.191	0.230	0.144	0.161	0.154	0.177	0.200
Panel B	NOINC vs. MONE-S		NOINC vs. MONE-L		NOINC vs. PROS-S		NOINC vs. PROS-L	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Complete'	Complete'	Complete'	Complete'	Complete'	Complete'	Complete'	Complete'
Treatment	−2.962	1.175	−0.428	4.510	0.149	0.814	1.505	−3.924
	(3.602)	(3.746)	(3.558)	(6.193)	(4.596)	(4.490)	(3.409)	(4.132)
Round2	−2.167	2.083**	−2.167	2.083**	−2.167	2.083**	−2.167	2.083**
	(3.081)	(1.096)	(3.029)	(1.062)	(3.074)	(1.072)	(2.982)	(1.046)
Treatment × Round2	5.125*	−0.353	9.952***	5.060***	7.333**	0.417	7.123**	7.197***
	(3.347)	(1.642)	(3.461)	(1.691)	(3.603)	(2.385)	(3.773)	(2.151)
High-productivity	N	Y	N	Y	N	Y	N	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
N. Obs	96	100	104	90	96	92	94	98
R-Square	0.268	0.210	0.225	0.128	0.124	0.208	0.188	0.187

Notes: $*$ p < 0.1; $**$ p < 0.05; $***$ p < 0.01. Dependent variable is the number of task completed. Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table compares the high productivity workers with the low productivity workers (based on their performance in Round 1) with their treatment effects. Panel A is for actual output, and Panel B is for total output. Standard errors cluster at the individual level.

Table A.8: Goal-Setting's effect between productivity levels

Panel A	(1)	(2)
	ΔComplete	ΔComplete
Round2 Goal	0.031	0.047*
	(0.044)	(0.036)
Round2 Goal × Achieve	0.161***	0.090***
	(0.049)	(0.018)
High-productivity	N	Y
Controls	Y	Y
N. Obs	122	118
R-Square	0.208	0.340
Panel B	(1)	(2)
	ΔComplete'	ΔComplete'
Round2 Goal	0.034	0.065**
	(0.047)	(0.037)
Round2 Goal × Achieve'	0.160***	0.093***
	(0.050)	(0.020)
High-productivity	N	Y
Controls	Y	Y
N. Obs	123	117
R-Square	0.200	0.323

Notes: $*$ p $<$ 0.1; $**$ p $<$ 0.05; $***$ p $<$ 0.01. Control variables include concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table display how Round 2 goals and its interaction with the goal-achievement affect participants' output improvement using the subgroup of low-productivity workers in column (1) and high-productivity workers in column (2). Panel A is for actual output, and Panel b is for total output. Standard errors cluster at the individual level.

Table A.9: Treatment effects between with and without self-chosen goals

Panel A	NOGOAL vs. NOINC		NG-PROS-S vs. PROS-S		NG-PROS-L vs. PROS-L	
	(1)	(2)	(3)	(4)	(5)	(6)
	Complete	Complete	Complete	Complete	Complete	Complete
With Goal Setting	12.108***	−5.960	8.354*	2.873	5.771	4.853
	(4.314)	(7.968)	(4.285)	(8.655)	(4.595)	(6.336)
Round2	1.184	1.043	3.854***	3.745***	3.833**	3.809**
	(1.795)	(1.799)	(1.361)	(1.323)	(1.610)	(1.649)
With Goal Setting × Round2	−0.413	−0.272	0.292	1.038	4.104*	4.129*
	(2.489)	(2.434)	(1.919)	(1.991)	(2.224)	(2.309)
Controls	N	Y	N	Y	N	Y
N. Obs	194	190	192	186	192	190
R-Square	0.059	0.168	0.038	0.202	0.044	0.217
Panel B	NOGOAL vs. NG-PROS-S		NOGOAL vs. NG-PROS-L		NG-PROS-S vs. NG-PROS-L	
	(1)	(2)	(3)	(4)	(5)	(6)
	Complete	Complete	Complete	Complete	Complete	Complete
NG-PROS-S	2.191	4.052				
	(4.423)	(5.297)				
NG-PROS-L			5.774	5.157	3.583	1.308
			(4.561)	(5.079)	(4.863)	(5.102)
Round2	1.184	1.043	1.184	1.043	3.854***	3.745***
	(1.732)	(1.838)	(1.732)	(1.851)	(1.335)	(1.376)
NG-PROS-S × Round2	2.670	2.702				
	(2.251)	(2.286)				
NG-PROS-L × Round2			2.650	2.766	−0.021	0.064
			(2.425)	(2.536)	(2.070)	(2.257)
Controls	N	Y	N	Y	N	Y
N. Obs	194	188	194	188	192	188
R-Square	0.008	0.170	0.023	0.113	0.011	0.186

Notes: ∗ p < 0.1; ∗∗ p < 0.05; ∗∗∗ p < 0.01. Dependent variable is the number of task completed correctly. Control variables include age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. Panel A compares between groups with and without self-chosen goals, and Panel B compares within groups without self-chosen goals. Standard errors cluster at the individual level.

Table A.10: Mediation Analysis of the Mechanism (Other Path 1)

	(1) Round2 Goal	(2) ΔComplete'	(3) ΔComplete'
Round2 Goal	NA	0.112***	0.065*
		(0.041)	(0.047)
Matched with society	11.444	7.599***	9.838***
	(11.264)	(3.140)	(3.222)
PROS-L	−2.313	7.459***	8.276***
	(9.241)	(3.151)	(3.274)
PROS-L × Matched with society	4.072	−8.782**	−15.339***
	(15.310)	(4.383)	(4.695)
(Personal) Matched			4.265
			(3.646)
PROS-L × (Personal) Matched			5.617
			(4.604)
Controls	Y	Y	Y
N. Obs	80	80	80
R-Square	0.276	0.346	0.428

Notes: $*$ p $<$ 0.1; $**$ p $<$ 0.05; $***$ p $<$ 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table includes participants who achieve their goals. Performance is valued by the total output. Standard errors cluster at the individual level.

Table A.11: Mediation Analysis of the Mechanism (Other Path 2)

Panel A	(1)	(2)	(3)
	Round2 Goal	ΔComplete	ΔComplete
Round2 Goal	NA	0.130***	0.100**
		(0.043)	(0.047)
Altruistic	8.694	−4.944	−4.549
	(11.466)	(3.540)	(3.702)
PROS-L	−7.883	1.181	1.032
	(9.796)	(3.097)	(3.373)
PROS-L × Altruistic	17.223	5.216	2.537
	(14.515)	(4.843)	(5.215)
(Personal) Matched			3.025
			(3.594)
PROS-L × (Personal) Matched			4.169
			(5.002)
Controls	Y	Y	Y
N. Obs	80	80	80
R-Square	0.329	0.326	0.377
Panel B	(1)	(2)	(3)
	Round2 Goal	ΔComplete'	ΔComplete'
Round2 Goal	NA	0.123***	0.098**
		(0.045)	(0.049)
Altruistic	8.694	−4.536	−4.145
	(11.466)	(3.749)	(3.886)
PROS-L	−7.883	1.240	1.024
	(9.796)	(3.200)	(3.498)
PROS-L × Altruistic	17.223	4.862	2.319
	(14.515)	(5.102)	(5.519)
(Personal) Matched			2.570
			(3.766)
PROS-L × (Personal) Matched			4.186
			(5.227)
Controls	Y	Y	Y
N. Obs	80	80	80
R-Square	0.329	0.316	0.358

Notes: $*$ p < 0.1; $**$ p < 0.05; $***$ p < 0.01. Control variables include the concern level, loss aversion, age, female (dummy), employment type (part-time, full-time, self-employed, other), and income level. This table includes participants who achieve their goals. Panel A is for actual output, and Panel B is for total output. Standard errors cluster at the individual level.

A.4 Some Screenshots of the Experiment

Mission match:

According to Charity Navigator (https://www.charitynavigator.org), there are five charities that are most in need. These are: Doctors Without Borders, The Nature Conservancy, UNICEF USA, DAV (Disabled American Veterans) Charitable Service Trust, and American Society for the Prevention of Cruelty to Animals. Please tell us what do you think about each of them?

Doctors Without Borders (https://www.doctorswithoutborders.org)

- Provides aid in nearly 60 countries to people whose survival is threatened by violence, neglect, or catastrophe, primarily due to armed conflict, epidemics, malnutrition, exclusion from health care, or natural disasters.

	Not Important	Slightly Important	Moderately Important	Important	Very Important
How much do you **personally** care about this issue?	○	○	○	○	○
How much do you think **society** cares about this issue?	○	○	○	○	○

The Nature Conservancy (https://www.nature.org)

- Address threats to conservation involving climate change, fire, fresh water, forests, invasive species, and marine ecosystems.

	Not Important	Slightly Important	Moderately Important	Important	Very Important
How much do you **personally** care about this issue?	○	○	○	○	○
How much do you think **society** cares about this issue?	○	○	○	○	○

UNICEF USA (https://www.unicefusa.org)

- Helps to save more children's lives than any other humanitarian organization, by providing health care and immunizations, clean water and sanitation, nutrition, education, emergency relief and more.

	Not Important	Slightly Important	Moderately Important	Important	Very Important
How much do you **personally** care about this issue?	○	○	○	○	○
How much do you think **society** cares about this issue?	○	○	○	○	○

DAV (Disabled American Veterans) Charitable Service Trust (https://www.dav.org)

- Provide direct service to ill, injured, or wounded veterans.

	Not Important	Slightly Important	Moderately Important	Important	Very Important
How much do you **personally** care about this issue?	○	○	○	○	○
How much do you think **society** cares about this issue?	○	○	○	○	○

American Society for the Prevention of Cruelty to Animals (https://www.aspca.org)

- Provides local and national leadership in animal-assisted therapy, animal behavior, animal poison control, anti-cruelty, humane education, legislative services, and shelter outreach.

	Not Important	Slightly Important	Moderately Important	Important	Very Important
How much do you **personally** care about this issue?	○	○	○	○	○
How much do you think **society** cares about this issue?	○	○	○	○	○

Instruction:

Part II: Doing Tasks (payments determined by your effort)

In the following experiment, you have the opportunity to earn money depending on your behavior. It is very important that you follow these rules.

The task consists of encoding combinations of letters (words) into numbers. In the task, three capital letters always yield a word. You have to allocate a number to each capital letter. The encryption code can be found in a table below the corresponding letter. Notice that five pairs of capital letters and numbers are provided to you each time, but only three pairs are needed for encoding.

V	D	R	Z	T
60	20	75	35	98

Let's take an example, You currently encrypt the word "**VRT**"

Correct Answer is: 607598

Notice: Your answer is always a six-digit number (Please do not add space or comma between any two-digit number. For example, 60 75 98 is a WRONG answer. And 607,598 is a WRONG answer, too.

Goal-setting:

Now the Real Task rounds begin. You are expected to finish two rounds of tasks, with 10 minutes in each round.

Answer a correct answer, you can win $0.02. You can do 200 tasks in each round at most.

Answer a wrong answer or Leave blank does NOT count finish one task and you cannot be paid in this task.

Now, you have a chance to set a goal for yourself.

My GOAL is to finish (enter an integer number ranges from 0 to 200) tasks:

Round 1 effort tasks:

Time Remaining: 09:10
You have set a goal to finish 100 tasks

TASK 3:

B	Y	K	T	E
33	26	98	12	17

You currently encrypt the word "**EYT**"

Write your answer in the below blank.

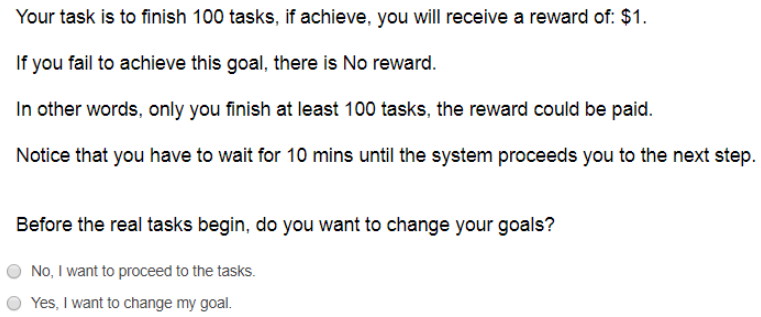
TASK 4:

S	D	A	Q	U
11	36	23	91	84

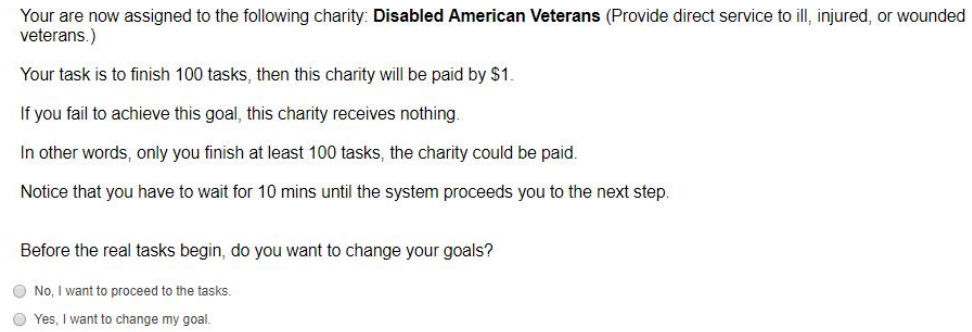
You currently encrypt the word "**QSU**"

Write your answer in the below blank.

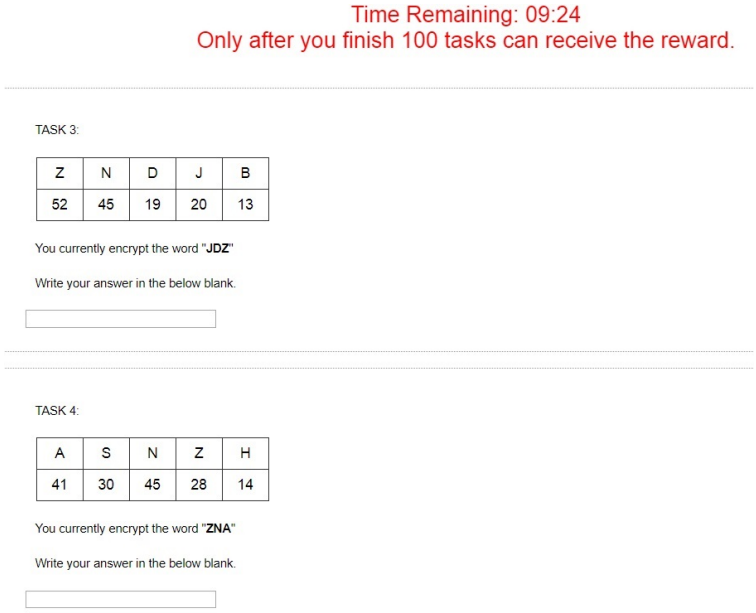
Round 2 monetary rewards (MONE-L):



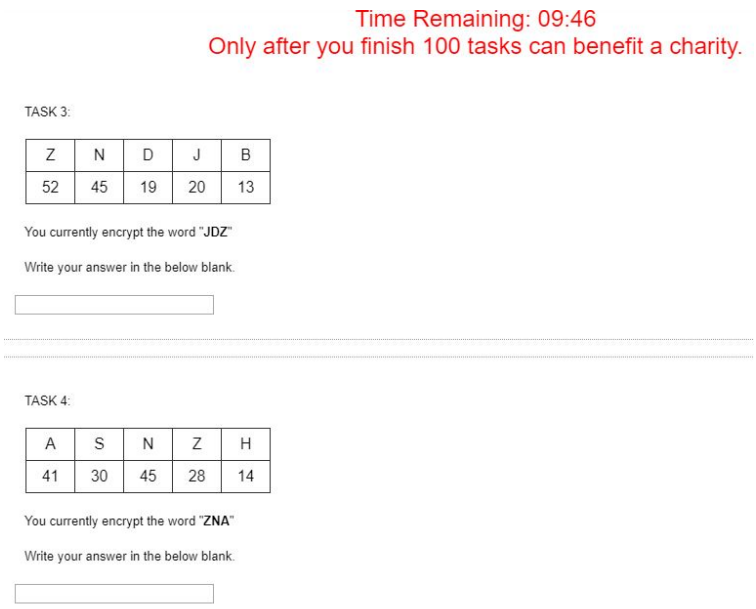
Round 2 assigning charities (PROS-L):



Round 2 effort tasks (monetary incentive):



Round 2 effort tasks (prosocial incentive):



B Chapter 2 Appendix

B.1 Full Model

Group 1

In Group 1 (G1), we consider a case where an individual i only chooses to donate effort (volunteer) now or not donate. A rational individual compares the utility of volunteering now with the utility of not volunteering. A typical individual's utility function at $t = n$ includes her endowments from now and the future, the effort cost of volunteering, the warm-glow effect, and the utility from knowing the charity beneficiaries receive the monetary donations. Hence,

If individual i in Group 1 **donates now**, her utility is:

$$\begin{aligned} u_{G1\,Donate\,Now}(e_{in}) =& u(W_{in}) - c(e_{in}) + a(G_{-i} + g_{it}) + v^e(g_{it}) + \beta_i\delta_i u(W_{if}) \\ =& lnW_{in} - c(e_{in}) + a_{1i}ln(G_{-i} + \hat{g_e}) + a^e_{2i}ln\hat{g_e} + \beta_i\delta_i ln(W_{if}) \end{aligned} \tag{10}$$

where $v^e(g_{it})$ denotes the warm-glow utility in effort donation, $a_{1i} \geq 0$ denotes the weight for pure altruism, and $a^e_{2i} \geq 0$ denotes the weight for warm-glow. The weight for pure altruism stands for the extent to which individual i cares about the total monetary donation to the charity, while the weight for warm-glow stands for the extent to which an individual i cares about her own donation (See Breman (2011)).

If individual i **rejects** to donate, her utility is:

$$u_{G1\,Reject}(0) = lnW_{in} + a_{1i}ln(G_{-i}) + \beta_i\delta_i ln(W_{if}) + r_i \tag{11}$$

where $r_i \leq 0$ denotes the disutility of rejecting as the individual feels pressured from not donating. We assume it is constant in both time periods (now and future).

Therefore, an individual i chooses to **donate now** if:

$$u_{G1\,Donate\,Now}(e_{in}) \geq u_{G1\,Reject}(0)$$

$$\leftrightarrow lnW_{in} - c(e_{in}) + a_{1i}ln(G_{-i} + \hat{g}_e) + a^e_{2i}ln\hat{g}_e + \beta_i\delta_i ln(W_{if}) \geq lnW_{in} + a_{1i}ln(G_{-i}) + \beta_i\delta_i ln(W_{if}) + r_i$$

$$ln[(\frac{G_{-i} + \hat{g}_e}{G_{-i}})^{a_{1i}}\, \hat{g}_e{}^{a^e_{2i}}] \geq c(e_{in}) + r_i \tag{12}$$

Group 2

In Group 2 (G2), an individual i chooses at $t = n$ between not volunteering (i.e., reject to donate) and volunteering at $t = f$. When the future arrives, she faces the choices between confirming or reneging on her promise to volunteer. We follow Andreoni and Serra-Garcia (2019)'s assumption that social pressure is experienced at the moment of making a decision to reject donation or renege on a promise to donate, while warm-glow is experienced at the moment of volunteering. Hence,

if individual i **rejects** to donate, her utility is:

$$u_{G2\,Reject}(0) = lnW_{in} + \beta_i\delta_i[a_{1i}ln(G_{-i}) + ln(W_{if})] + r_i \tag{13}$$

if individual i **plans to donate** at $t = n$ and **confirms donation** at $t = f$, her utility is:

$$\begin{aligned} u_{G2\,Confirm}(e_{if}) =& u(W_{in}) + \beta_i\delta_i[v^e(\hat{g}_e) - c(e_{if}) + a(G_{-i} + \hat{g}_e) + u(W_{if})] \\ =& lnW_{in} + \beta_i\delta_i[a^e_{2i}ln\hat{g}_e - c(e_{if}) + a_{1i}ln(G_{-i} + \hat{g}_e) + ln(W_{if})] \end{aligned} \tag{14}$$

if individual i **plans to donate** at $t = n$ but **reneges** at $t = f$, her utility is:

$$u_{G2\,Renege}(0, L_i) = lnW_{in} + \beta_i\delta_i[a_{1i}ln(G_{-i}) + ln(W_{if}) + L_i] \tag{15}$$

where L_i represents the disutility of reneging with $L_i \leq 0$.

Therefore, an individual i in Group 2 **plans to donate** at $t = n$ and chooses to **donate** over **reneging** at $t = f$ if at $t = f$:

$$u_{G2\,Confirm}(e_{if}) \geq u_{G2\,Renege}(0, L_i)$$

$$\leftrightarrow [a_{2i}^{e} ln\hat{g}_e - c(e_{in}) - \epsilon_{if} - \xi_{\epsilon_{if}} + a_{1i} ln(G_{-i} + \hat{g}_e) + ln(W_{if})] \geq [a_{1i} ln(G_{-i}) + ln(W_{if}) + L_i + \xi_{L_i}]$$

$$ln[(\frac{G_{-i} + \hat{g}_e}{G_{-i}})^{a_{1i}}\, \hat{g}_e{}^{a_{2i}^{e}}] \geq c(e_{in}) + \epsilon_{if} + L_i + \xi_{\epsilon_{if}} + \xi_{L_i} \tag{16}$$

Given that $c(e_{if})$ and L_i are the only two unknowns for individual i at $t = n$, we assume $L_i \sim U[-l, 0]$ and $\epsilon_{if} \sim U[-e, e]$. Let's assume their PDFs are $f(\epsilon_{if})$, $f(L_i)$, their CDFs are $F(\epsilon_{if})$, $F(L_i)$, and they are continuous and strictly increasing over the assumed domains. $\xi_{\epsilon_{if}}$ and ξ_{L_i} stand for the shocks with equal means, which equals to the difference between the actual ϵ_{if} and L_i and the predicted. Let's say $\gamma_i = ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\, \hat{g}_e{}^{a_{2i}^{e}}] - c(e_{in}) - \xi_{\epsilon_{if}} - \xi_{L_i}$, by which we can derive the probability of confirming or reneging in the future (at $t = f$). If $\epsilon_{if} + L_i \leq \gamma_i$ then an individual's prior is to confirm effort donation, and if $\epsilon_{if} + L_i \geq \gamma_i$, then this individual's prior is to renege. In other words, expected higher participation cost and/or lower reneging cost results in a higher prior probability of reneging.[62] Therefore, after knowing one's own prior probability of reneging (we call it M_i), an individual i prefers to **plan to donate** at $t = n$ to **reject** to donate if:

$$\mathbb{E}\, u_{G2\,Plan\,to\,Donate}(e_{if}) \geq u_{G2\,Reject}(0)$$

$$\leftrightarrow \mathbb{E}[Pr(Renege_i)\; u_{G2\,Renege}(0, L_i)] + \mathbb{E}[Pr(Confirm_i)\; u_{G2\,Confirm}(e_{if})] \geq u_{G2\,Reject}(0)$$

$$ln[(\frac{G_{-i} + \hat{g}_e}{G_{-i}})^{a_{1i}}\, \hat{g}_e{}^{a_{2i}^{e}}] \geq \frac{\frac{r_i}{\beta_i \delta_i} - M_i L_i}{1 - M_i} + c(e_{in}) + \epsilon_{if} \tag{17}$$

In summary, given her prior belief on the probability of reneging in the future, an individual does not reject to volunteer if (8) is satisfied. On the contrary, if equation (8) is not

[62] We discuss the ranges of reneging cost (L_i) and participation cost (ϵ_{if}) in Appendix B.2.

satisfied, this individual prefers to reject volunteering directly. When the future arrives, if equation (7) is satisfied, she will confirm her effort donation; otherwise, she will renege on her promise to volunteer.

Group 3

In Group 3 (G3), we consider a case where an individual i can choose among three options: donate now, donate on a future date, or reject to donate time and effort. When the future date arrives, she has the choice to confirm her promised donation or renege on her promise.

Therefore, as in Group 2, for an individual i who **plans to donate** at $t = n$ and **confirms donation** at $t = f$, equations (7) and (8) need to be satisfied. However, for an individual i who **plans to donate** at $t = n$ but **reneges** at $t = f$, equation (8) but not equation (7) needs to be satisfied.

At $t = n$, an individual i who prefers to **donate now** requires the utility of donating now to be larger than both the utility of **planning to donate** and the utility of **rejecting** to donate. After some manipulation, we obtain the following inequalities:

$$(i)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}] \ \geq\ \frac{1}{1-\beta_i\delta_i+\beta_i\delta_i M_i}[\beta_i\delta_i(M_i-1)\epsilon_{if}+\beta_i\delta_i M_i L_i+ (1-\beta_i\delta_i+\beta_i\delta_i M_i)c(e_{in})+(\beta_i\delta_i-1)a_{1i}lnG_{-i}] \qquad (A)$$

$$(ii)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}] \ \geq\ c(e_{in})+r_i \qquad (C)$$

We call the RHS of expressions (i) and (ii) above A and C, respectively. If the individual prefers **planning to donate** to the other two options, the following expressions need to be satisfied:

$$(iii)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}] \ \leq\ \frac{1}{1-\beta_i\delta_i+\beta_i\delta_i M_i}[\beta_i\delta_i(M_i-1)\epsilon_{if}+\beta_i\delta_i M_i L_i+ (1-\beta_i\delta_i+\beta_i\delta_i M_i)c(e_{in})+(\beta_i\delta_i-1)a_{1i}lnG_{-i}] \qquad (A)$$

$$(iv)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}]\ \ \geq\ \ \frac{\frac{r_i}{\beta_i\delta_i}-M_iL_i}{1-M_i}+c(e_{in})+\epsilon_{if}\qquad (B)$$

Notice that the RHS in expression (iii) is the same as the RHS in (i). We call the RHS of (iv) B. Further, if this individual chooses to **reject** to donate rather than the other two options, the following conditions should be satisfied:

$$(v)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}]\ \ \leq\ \ \frac{\frac{r_i}{\beta_i\delta_i}-M_iL_i}{1-M_i}+c(e_{in})+\epsilon_{if}\qquad (B)$$

$$(vi)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}]\ \ \leq\ \ c(e_{in})+r_i\qquad (C)$$

Finally, when the future date arrives, for an individual who prefers **donating** to **reneging**:

$$u_{G3\,Confirm}(e_{if})\ \ \geq\ \ u_{G3\,Renege}(0,L_i)$$

$$(vii)\ \ ln[(\frac{G_{-i}+\hat{g}_e}{G_{-i}})^{a_{1i}}\ \hat{g}_e{}^{a^e_{2i}}]\ \ \geq\ \ c(e_{in})+\epsilon_{if}+L_i+\xi_{\epsilon_{if}}+\xi_{L_i}\qquad (D)$$

We call the RHS of (vii) (D).

Notice that in expressions $(i)-(vii)$, the LHS expressions are exactly the same which we call λ_i. This expression is a function of pure altruism and warm-glow. We can use this expression and the inequities above to establish relationships between A, B, and C. Indeed, there are six possible relationships between A, B, and C with the only reasonable ones being: 1) $A>B>C$, 2) $A>C>B$, and 3) $C>A>B$. In Appendix B.2 we prove that only 2) $A>C>B$ could exist.

B.2 Mathematical Proofs

Priors:

Given the assumption of $L_i \sim U[-l, 0]$ and $\epsilon_{if} \sim U[-e, e]$, we set $z_i = L_i + \epsilon_{if}$, and write L_i as $z_i - \epsilon_{if}$, such that $-l \leq z_i - \epsilon_{if} \leq 0$, which is equivalent to $z_i \leq \epsilon_{if} \leq l + z_i$. We will prove the ranges for ϵ and l can determine participant's prior probability of reneging as follows.

Case (i): If $l \leq 2e$ (which is equivalent to $-e \leq e - l$):

$$\begin{cases} (1)\text{If} - e - l \leq z \leq -e, \text{then} - e \leq \epsilon \leq l + z; \\ (2)\text{If} - e \leq z \leq e - l, \text{then } z \leq \epsilon \leq l + z \\ (3)\text{If } e - l \leq z \leq e, \text{then } z \leq \epsilon \leq e. \end{cases}$$

Now, given that $f(\epsilon) = \frac{1}{2e}$, $f(L) = \frac{1}{l}$, and $0 \leq l \leq 2e$, we have:

(1) If $-e - l \leq z \leq -e$, $f_\gamma(z) = \frac{l+z+e}{2e}$, $F_\gamma(z) = \frac{(l+z+e)^2}{2el}$

Then, the minimum of $F_\gamma(z)$ exists when $z = -e-l$, at which $F_\gamma(z) = \Pr(\gamma \leq z) = 0 \in [0, 1]$ always holds. In addition, the maximum of $F_\gamma(z)$ exists when $z = -e$, at which $\Pr(\gamma \leq z) = \frac{l}{2e} \in [0, 1]$.

(2) If $-e \leq z \leq e - l$, $f_\gamma(z) = \frac{1}{2e}$, $F_\gamma(z) = \frac{l}{2e} + \frac{z+e}{2e} = \frac{z+e+l}{2e}$;

Then, the minimum of $F_\gamma(z)$ exists when $z = -e$, at which $F_\gamma(z) = \Pr(\gamma \leq z) = \frac{l}{2e} \in [0, 1]$ always holds. The maximum of $F_\gamma(z)$ exists when $z = e - l$, at which $\Pr(\gamma \leq z) = 1 \in [0, 1]$.

(3) If $e - l \leq z \leq e$, $f_\gamma(z) = \frac{e-z}{2el}$, $F_\gamma(z) = \frac{l}{2e} + \frac{2e-l}{2e} + \frac{(e-z)(z-e+l)}{2el}$.

Then, the minimum of $F_\gamma(z)$ exists when $z = e - l$ or $z = e$, at which $F_\gamma(z) = \Pr(\gamma \leq z) = 1$ always holds. That is, $\forall e - l \leq z \leq e$, $F_\gamma(z) = 1$, so that the expected probability of reneging in the future is equal to 1.

Case (ii): If $l \geq 2e$ (which is equivalent to $e - l \leq -e$):

$$\begin{cases} (1)\text{If} -e-l \le z \le e-l, \text{then} -e \le \epsilon \le l+z; \\ (2)\text{If } e-l \le z \le -e, \text{then} -e \le \epsilon \le e; \\ (3)\text{If} -e \le z \le e, \text{then } z \le \epsilon \le e. \end{cases}$$

Now, given that $f(\epsilon) = \frac{1}{2e}$, $f(L) = \frac{1}{l}$, and $l \ge 2e$, we have:

(1) If $-e-l \le z \le e-l$, $f_\gamma(z) = \frac{l+z+e}{2el}$, $F_\gamma(z) = \frac{(l+z+e)^2}{2el}$

Here, the minimum of $F_\gamma(z)$ exists when $z = -e-l$, at which $F_\gamma(z) = \Pr(\gamma \le z) = 0 \in [0,1]$ always holds. In addition, the maximum of $F_\gamma(z)$ exists when $z = e-l$, at which $\Pr(\gamma \le z) = \frac{2e}{l} \in [0,1]$.

(2) If $e-l \le z \le -e$, $f_\gamma(z) = \frac{1}{l}$, $F_\gamma(z) = \frac{2e}{l} + \frac{Z-e+l}{l} = \frac{z+e+l}{l}$

Here, the minimum of $F_\gamma(z)$ exists when $z = e-l$, at which $F_\gamma(z) = \Pr(\gamma \le z) = \frac{2e}{l} \in [0,1]$ always holds. The maximum of $F_\gamma(z)$ exists when $z = -e$, at which $\Pr(\gamma \le z) = 1 \in [0,1]$.

(3) If $-e \le z \le e$, $f_\gamma(z) = \frac{e-z}{2el}$, $F_\gamma(z) = \frac{2e}{l} + \frac{l-2e}{l} + \frac{e^2-z^2}{2el} = 1 + \frac{e^2-z^2}{2el}$

Here, the minimum of $F_\gamma(z)$ exists when $z = -e$ or $z = e$, at which $F_\gamma(z) = \Pr(\gamma \le z) = 1$ always holds. That is, $\forall -e \le z \le e$, $F_\gamma(z) = 1$, so that the expected probability of reneging in the future is equal to 1.

In conclusion, there always exists a threshold for z such that in case (i) $z > e-l$ and in case (ii) $z > -e$, an individual's prior probability of reneging in the future equals to 1. Furthermore, case (i) suggests that when reneging cost increases and/or participation cost decreases, the prior probability of reneging increases. However, case (ii) suggests differently: if reneging cost decreases and/or participation cost increases, the prior probability of reneging increases. We expect the prior increases with $\epsilon_{if} + L_i$, so case (ii) fits our settings better.

Critical Values:

Prove only $A > C > B$ exists: (Given $A > B$ from above three relationship)

Set $\frac{(\beta_i\delta_i - 1)}{1-\beta_i\delta_i+\beta_i\delta_i M_i} a_{i1} lnG_{-i} = \text{s}$, where $\text{s} < 0$ is known from the model.

(1) $\forall A > B$, we have $C > B$:

$$A > B \Leftrightarrow A - B > 0$$

$$\Leftrightarrow -\frac{1}{1-\beta_i\delta_i+\beta_i\delta_i M_i}\epsilon_{if} - \frac{1}{\beta_i\delta_i(1-M_i)}r_i + [\frac{M_i\beta_i\delta_i}{1-\beta_i\delta_i+\beta_i\delta_i M_i} + \frac{M_i\beta_i\delta_i}{\beta_i\delta_i(1-M_i)}]L_i + \frac{\beta_i\delta_i - 1}{\beta_i\delta_i}c(e_{i1}) + s > 0$$

Set $\beta_i\delta_i - \beta_i\delta_i M_i = x$, $0 < x < 1$

$$\Leftrightarrow -\frac{1}{x}r_i + [\frac{M_i\beta_i\delta_i}{1-x} + \frac{M_i\beta_i\delta_i}{x}]L_i + s > \frac{\epsilon_{if}}{1-x}$$

$$\Leftrightarrow -\frac{1}{x}r_i + \frac{M_i\beta_i\delta_i}{(1-x)x}L_i + s > \frac{\epsilon_{if}}{1-x}$$

$$\Leftrightarrow -\frac{1-x}{x}r_i + \frac{M_i}{1-M_i}L_i - \epsilon_{if} + s(1-x) > 0$$

$$\Leftrightarrow (\frac{\beta_i\delta_i - \beta_i\delta_i M_i - 1}{\beta_i\delta_i - \beta_i\delta_i M_i})r_i + \frac{M_i}{1-M_i}L_i - \epsilon_{if} > 0 \quad (\Delta)$$

$$or \Leftrightarrow (\beta_i\delta_i - \beta_i\delta_i M_i - 1)r_i + M_i\beta_i\delta_i L_i - (\beta_i\delta_i - \beta_i\delta_i M_i)\epsilon_{if} + s(1-x) > 0 \quad (\Delta\Delta)$$

$$\begin{aligned} C - B &= -\epsilon_{if} + [\frac{\beta_i\delta_i - \beta_i\delta_i M_i}{\beta_i\delta_i(1-M_i)} - \frac{1}{\beta_i\delta_i(1-M_i)}]r_i + \frac{M_i}{1-M_i}L_i \\ &= -\epsilon_{if} + [\frac{\beta_i\delta_i - \beta_i\delta_i M_i - 1}{\beta_i\delta_i(1-M_i)}]r_i + \frac{M_i}{1-M_i}L_i \\ &= (\Delta) \\ &\Rightarrow\ C > B \end{aligned}$$

(2) $\forall A > B$, we have $A > C$:

$$A - C = \frac{\beta_i\delta_i - \beta_i\delta_i M}{1-\beta_i\delta_i+\beta_i\delta_i M}\epsilon_{if} - r_i + \frac{M\beta_i\delta_i}{1-\beta_i\delta_i+\beta_i\delta_i M}L_i + s$$

Set $\beta_i\delta_i - \beta_i\delta_i M = x$, $0 < x < 1$

$$
\begin{aligned}
A - C &= \frac{x}{1-x}\epsilon_{if} - \frac{1-x}{1-x}r_i + \frac{M_i\beta_i\delta_i}{1-x}L_i + s \\
&= \frac{1}{1-x}[(\beta_i\delta_i - \beta_i\delta_i M - 1)r_i + M_i\beta_i\delta_i L_i - (\beta_i\delta_i - \beta_i\delta_i M_i)\epsilon_{if} + s(1-x)] \\
&= \frac{1}{1-x}(\Delta\Delta) \\
&\Rightarrow\ A > C \\
&\Rightarrow\ A > C > B
\end{aligned}
$$

Warm-Glow Effects:

In our experimental settings, the expected individual participation costs and reneging costs across the three volunteering treatments and the monetary treatment, as well as their shocks are assumed to be the same:

$$
\begin{cases}
Pr(Renege_e) = M_i^e = Pr(\gamma_i(e) \le \epsilon_{if}^e + L_i^e) = Pr(\gamma_i(e) \le \epsilon_{if} + L_i) \\
\\
Pr(Renege_m) = M_i^m = Pr(\gamma_i(m) \le \epsilon_{if}^m + L_i^m) = Pr(\gamma_i(m) \le \epsilon_{if} + L_i)
\end{cases}
$$

$$
\begin{aligned}
&s.t.\ Pr(Renege_e)] \le Pr(Renege_m)] \\
&\quad \Leftrightarrow \gamma_i(e) \ge \gamma_i(m) \\
&\quad \Leftrightarrow ln[(\frac{G_{-i} + \hat{g_e}}{G_{-i}})^{a_{1i}}\ \hat{g_e}^{a_{2i}^e}] - c(e_{in}) \ge ln[(\frac{G_{-i} + \hat{g_e}}{G_{-i}})^{a_{1i}}\ \hat{g_e}^{a_{2i}^m}] - c(\hat{g'}, m) \\
&\quad \Leftrightarrow a_{2i}^e - a_{2i}^m \ge \frac{c(e_{in}) - c(\hat{g_m}, s)}{ln\hat{g_e}}
\end{aligned}
\tag{18}
$$

Therefore, for an individual i in G2: if equation (9) is satisfied, he or she has prior as Pr(reneging in effort donation) $\le$ Pr(reneging in monetary donation). In addition, in our experimental setting, we assume $c(e_{in}) = c(\hat{g_m}, s) = \$1$. As a result, equation (9) is always satisfied in our case. That is, effort donation is more motivated by the warm-glow effect than the monetary donation. This is consistent with Lilley and Slonim (2014) and Brown et al. (2018)'s findings. Thus, in G2, we can assume an individual has $a_{2i}^e - a_{2i}^m = u_i \ge 0$, at which $M_i^m - M_i^e = \mu_i \ge 0$, which is equivalent to say this individual has prior of Pr(reneging

in effort donation) $\leq$ Pr(reneging in monetary donation).

$$\begin{cases} A_e = \frac{1}{1-\beta_i\delta_i+\beta_i\delta_i M_i^e}[\beta_i\delta_i(M_i^e-1)\epsilon_{if} + \beta_i\delta_i M_i^e L_i + (1-\beta_i\delta_i+\beta_i\delta_i M_i^e)c(e_{in}) + (\beta_i\delta_i-1)a_{1i}lnG_{-i}] \\ A_m = \frac{1}{1-\beta_i\delta_i+\beta_i\delta_i M_i^m}[\beta_i\delta_i(M_i^m-1)\epsilon_{if} + \beta_i\delta_i M_i^m L_i + (1-\beta_i\delta_i+\beta_i\delta_i M_i^m)c(\hat{g_m}, s) + (\beta_i\delta_i-1)a_{1i}lnG_{-i}] \end{cases}$$

$$\begin{cases} B_e = \frac{\frac{r_i^e}{\beta_i\delta_i}-M_i^e L_i}{1-M_i^e} + c(e_{in}) + \epsilon_{if} \\ B_m = \frac{\frac{r_i^m}{\beta_i\delta_i}-M_i^m L_i}{1-M_i^m} + c(\hat{g_m}, s) + \epsilon_{if} \end{cases}$$

$$\begin{cases} D'_e = \beta_i\delta_i[c(e_{in}) + L_i + \epsilon_{if} + \xi_{\epsilon_{if}} + \xi_{L_i}] \\ D'_m = \beta_i\delta_i[c(\hat{g_m}, s) + L_i + \epsilon_{if} + \xi_{\epsilon_{if}} + \xi_{L_i}] \end{cases}$$

Based on above assumptions, we have $D'_e = D'_m$ and $\bar{\lambda}_e \geq \bar{\lambda}_m$. For simplicity, we normalize $\beta_i\delta_i = 1$ [63], and let shocks equal to zero. Now, the following equations should hold:

$$\begin{aligned} D'_e = D'_m &= L_i + c(e_{in}) + \epsilon_{if} = L_i + c(\hat{g_m}, s) + \epsilon_{if} \\ B_e - B_m &= \frac{\mu_i(L_i - r_i^m) + (1-M_i^e-\mu_i)(r_i^e - r_i^m)}{(1-M_i^e)(1-M_i^e-\mu_i)} \end{aligned} \tag{19}$$

Hence, to compare reneging rates in effort donation and monetary donation treatments, i.e., $\frac{D'_e - B_e}{\bar{\lambda}_e - B_e}$ and $\frac{D'_m - B_m}{\bar{\lambda}_m - B_m}$, it is equivalent to solve the equation:

$$(D'_e - B_e)(\bar{\lambda}_m - B_m) - (D'_m - B_m)(\bar{\lambda}_e - B_e) \tag{20}$$

For G2, since $\bar{\lambda}_e \geq \bar{\lambda}_m \geq D'_e = D'_m$, we can rewrite equation (11) as:

$$\begin{aligned} &(\bar{\lambda}_e - \bar{\lambda}_m)(B_m - D'_m) + (\bar{\lambda}_m - D'_m)(B_m - B_e) \\ \leftrightarrow\ &(\bar{\lambda}_e - \bar{\lambda}_m)\frac{r_i^m - L_i}{1-M_i^e-\mu_i} + (\bar{\lambda}_m - D'_m)\frac{\mu_i(r_i^m - L_i) + (1-M_i^e-\mu_i)(r_i^m - r_i^e)}{(1-M_i^e)(1-M_i^e-\mu_i)} \end{aligned} \tag{21}$$

Based on results in section 2.3 where we find evidence of significantly larger than zero

[63] This is a reasonable assumption under a 4-day time period. We test individuals' time preference using a quantitative time preference survey (Falk et al., 2016), and the average discount rate over a 4-day period is 0.996 for both Group 2 and Group 2M.

reneging rate in G2. This means $B_e < D'_e$ is always held, so rejection cost is always larger than the reneging cost in volunteering treatment, i.e., $L_i \geq r_i^e$. Now we consider the following three possible cases: $r_i^m \geq L_i \geq r_i^e$, $L_i \geq r_i^m \geq r_i^e$, and $L_i \geq r_i^e \geq r_i^m$. We can reasonably assume that the cost of reneging on a promise to donate money is higher than the cost of rejecting to donate money. In absolute value $|r_i^m| \leq |L_i^m|$. We can also reasonably assume that $|L_i^e| \leq |r_i^e|$. Our experimental data suggests that the average value of altruism index for individuals who plan to donate equals to 0.890 in G2M treatment, which is significantly higher than that in the volunteering treatment (G2), where the average altruism index equals to 0.457 (χ^2 = *1.81, p = 0.091*). However, the overall altruism index are indifferent with each other for two groups, where the average altruism index equals to 0.097 in G2 and -0.162 in G2M (χ^2 = *2.37, p = 0.125*). These findings lead to our conclusion of $B_e < B_M$, which is equivalent to say $r_i^e < r_i^m$ is held. In other words, for individuals who plan to donate, lower altruism index in effort giving than in monetary donation means rejection cost in effort giving is greater than that in monetary giving. Furthermore, we prove $B_e < B_M$ and $\bar{\lambda_e} < \bar{\lambda_M}$, so we can predict the rate of planning to donate in G2 is higher than that in the monetary treatment (G2M). Our results support this finding as well (χ^2 = *21.14, p = 0.000*). Therefore, we prove that $L_i \geq r_i^e \geq r_i^m$ does not exist in our experiment.

If $r_i^m \geq L_i \geq r_i^e$ is held, this means for an individual in the monetary treatment who delays rejection, even if reneging costs are higher than rejection costs, will only do so if she has higher altruism index and/or experiences higher warm-glow effect than in the volunteering treatment. In addition, with larger reneging costs than rejection costs in monetary donation, those who are willing to donate on a future data due to "social pressure" are crowded out and only those who really want to donate plan to donate. Therefore, we predict that the reneging rate for volunteering should be higher than for the monetary treatment where the reneging rate is indifferent from zero. However, if $L_i \geq r_i^m \geq r_i^e$ is held, the relationship of reneging rates between the two treatments is ambiguous, but this case exists only if the reneging rates in both treatments are significantly larger than zero.

B.3 Altruism

To see the effect of altruism on volunteering, we constructed an altruism index using Principal Components Analysis (PCA). The index is a representative measure of altruism that includes both qualitative and quantitative questions. The qualitative questions included attitudes towards charities and pro-social behaviors. The quantitative question was the amount sent to the charity in a dictator game. We used the first component of the principal component analysis (PCA) as our altruism index measure, which explained 62.94% of the total variance.

Table B.1 summarizes the results of regressing the probability of donating effort on the altruism index and a list of controls including sex, age, attending college or not, being white or not, having higher than average IQ score or not, confidence about one's IQ, having religious belief or not, number of children, and number of siblings. Using a logit regression specification, we found that one standard-deviation increase in altruism index significantly increased the probability of donating effort by 9.9, 7.9, and 8.0 percentage points in G1, G2, and G3, respectively.

In addition, we found that individuals' altruism index was negatively related to the reneging rate in G2. A one standard-deviation increase in the altruism index significantly decreased the probability of reneging by 8.2 percentage points. In contrast, altruism index was positively related to the reneging rate in G3. A one standard-deviation increase in the altruism index significantly increased the probability of reneging by 15.6 percentage points. The regression results are presented in Table B.2.[64]

We see that the regression results are aligned with our model predictions. The model predicts lower reneging rates in G2 and higher reneging rates in G3 when the altruism level increases. When altruism increases, those who can donate at $t = n$ will do so rather than wait until the future date. This is represented in Figure B.4 by A moving to the left and $\bar{\lambda}$ moving to the right, whereas both B and D' are unchanged. That is, more people donate at $t = n$ in G3 while more people plan to donate in G2. However, because the proportion

[64]Using step-wise regressions for column (4), (5) and (6), we find the results are consistent.

Table B.1: Effects of Altruism on Volunteering

	(1)	(2)	(3)	(4)	(5)	(6)
Altruism Index (G1)	0.111***			0.099***		
	(0.031)			(0.035)		
Altruism Index (G2)		0.110***			0.079***	
		(0.020)			(0.023)	
Altruism Index (G3)			0.085***			0.080***
			(0.022)			(0.022)
Controls	No	No	No	Yes	Yes	Yes
Obs	134	178	228	134	178	228
Pseudo R^2	0.071	0.095	0.043	0.154	0.180	0.093

Note: Coefficients are marginal effects. Control variables are gender (dummy), age, college (dummy), white (dummy), high IQ score (dummy), confidence about one's IQ, have religious belief (dummy), number of children, and number of siblings. $*$ Significant at the 10% level; $**$ Significant at the 5% level; $***$ Significant at the 1% level.

Table B.2: Effects of Altruism on Reneging

	(1)	(2)	(3)	(4)
Altruism Index (G2)	−0.068**		−0.082**	
	(0.025)		(0.022)	
Altruism Index (G3)		0.061		0.156**
		(0.052)		(0.073)
Controls	No	No	Yes	Yes
Obs	67	33	67	33
Pseudo R^2	0.133	0.027	0.559	0.346

Note: Coefficients are marginal effects. Control variables are gender (dummy), age, college (dummy), white (dummy), high IQ score (dummy), confidence about one's IQ, have religious belief (dummy), number of children, and number of siblings. $*$ Significant at the 10% level; $**$ Significant at the 5% level; $***$ Significant at the 1% level.

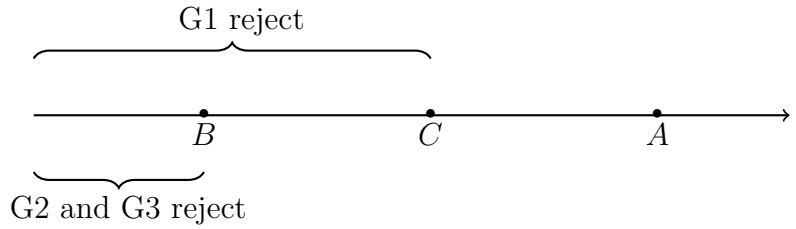
of those who renege is unchanged in both groups, the ratio of reneging to planned donation goes up for G3 and goes down for G2.

B.4 Figures

1. Rejection:

- On Day 1, G1 has higher rejection rates than G2 and G3.

FIGURE B.1

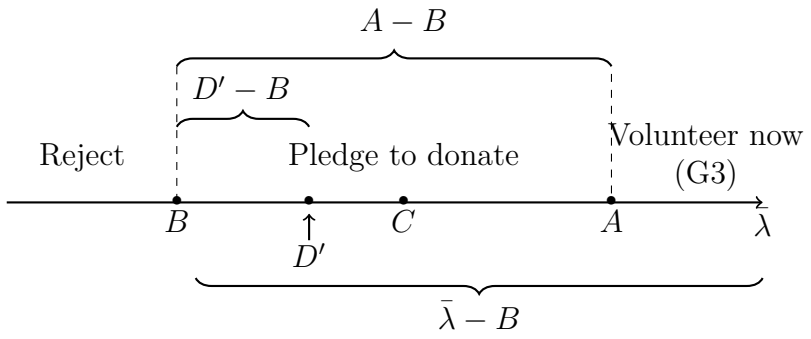


2. Reneging:

- G3 reneging rate $= \dfrac{D' - B}{A - B}$, G2 reneging rate $= \dfrac{D' - B}{\bar{\lambda} - B}$ where $\bar{\lambda}$ is the maximum possible value of λ

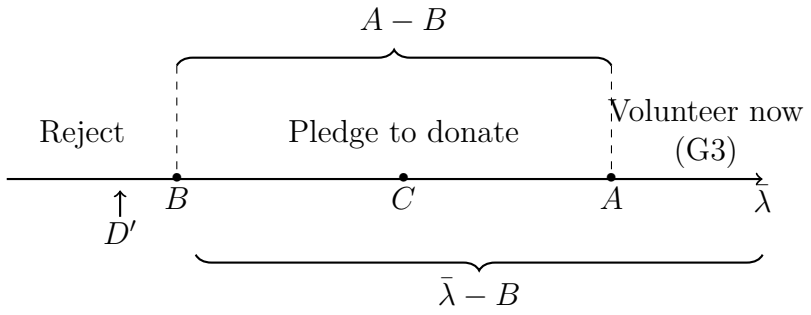
- When $\bar{\lambda} > D' > B$, G3 reneging rate > G2 reneging rate

FIGURE B.2.1



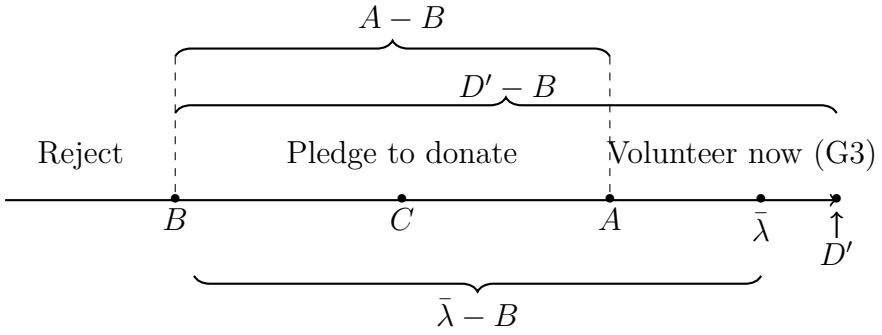
- When $D' < B$, G3 reneging rate = G2 reneging rate (no one reneges)

FIGURE B.2.2



- When $D' > \bar{\lambda}$, G3 reneging rate = G2 reneging rate (everyone reneges)

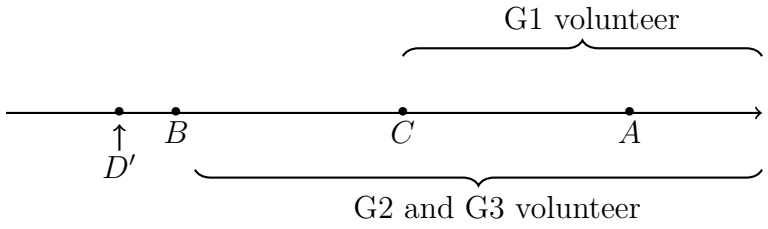
FIGURE B.2.3



3. Volunteering:

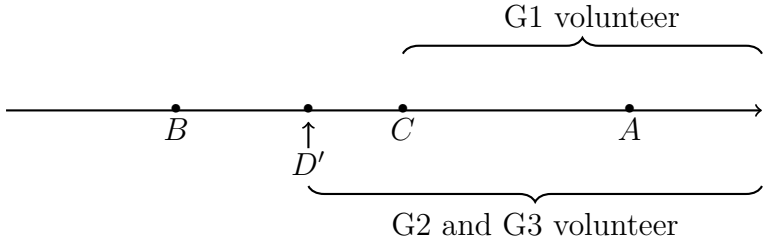
- When $C \geq B \geq D'$, $V_{G2} = V_{G3} \geq V_{G1}$

FIGURE B.3.1



- When $C \geq D' \geq B$, $V_{G2} = V_{G3} \geq V_{G1}$

FIGURE B.3.2



- When $D' \geq C \geq B$, $V_{G1} \geq V_{G3} \geq V_{G2}$

FIGURE B.3.3

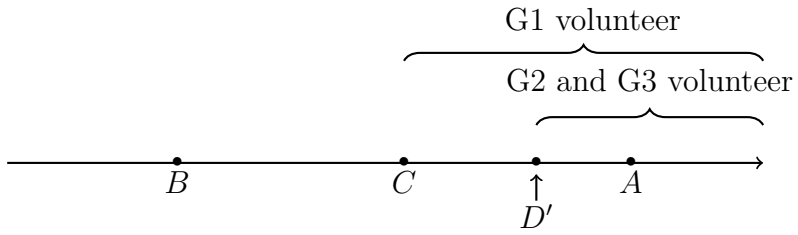
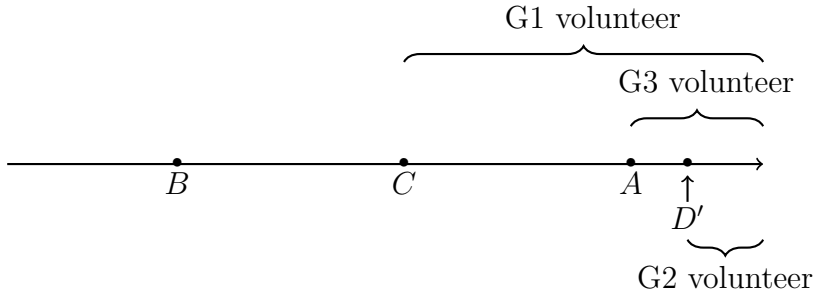


FIGURE B.3.4



4. Altruism:

- More altruism leads to higher overall donations. However, more altruism decreases reneging rate in G2 and increases reneging rate in G3.

FIGURE B.4

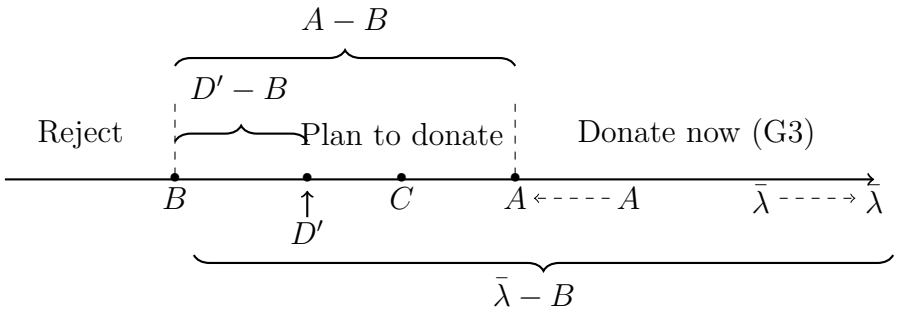
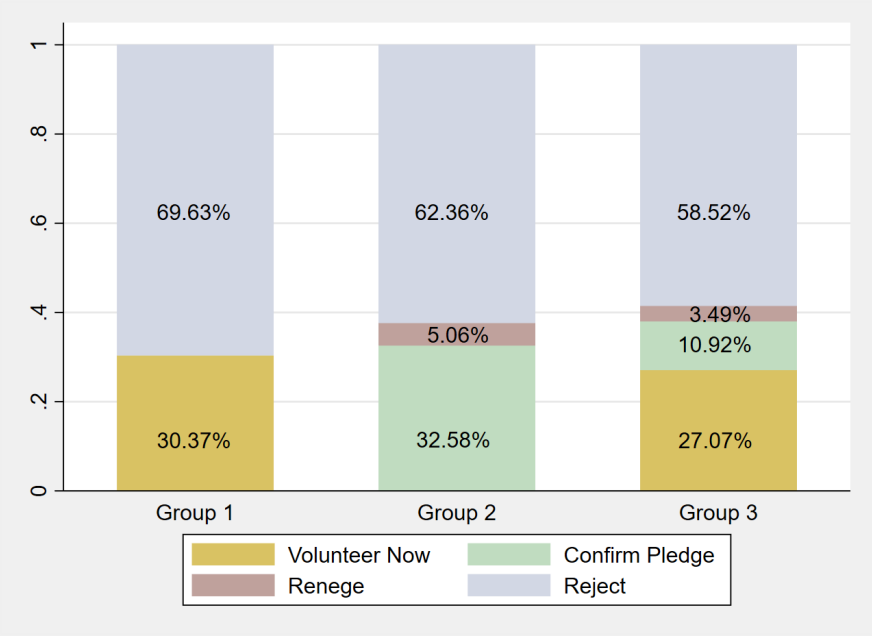


FIGURE B.5


0
.2
.4
.6
.8
1
69.63%
30.37%
62.36%
5.06%
32.58%
58.52%
3.49%
10.92%
27.07%
Group 1
Group 2
Group 3
Volunteer Now
Confirm Pledge
Renege
Reject

B.5 Tables

Table B.3: Reneging Rates

	(1)	(2)
G3	0.108*	0.115*
	(0.086)	(0.088)
Controls	No	Yes
Obs	100	100
Pseudo R^2	0.019	0.175

Note: Coefficients are marginal effects. Control variables are gender (dummy), age, college (dummy), white (dummy), high IQ score (dummy), confidence about one's IQ, have religious belief (dummy), number of children, number of siblings, and altruism index. $*$ Significant at the 10% level; $**$ Significant at the 5% level; $***$ Significant at the 1% level.

C Chapter 3 Appendix

C.1 Robustness Checks

Reconstruct the Index:

We constructed our index only using the qualitative questions, so it is possible the index could be different if we construct the index including the quantitative questions as well. We define this new index with both qualitative questions and quantitative questions as "index2".

The first quantitative question sets a situation as: *You won 1,000 Euro in a lottery. Considering your current situation, how much would you donate to charity? (Values between 0 and 1,000 are allowed).* For this question, we separated participants by their altruism level in this question (high altruism and low altruism with 128 and 124 individuals, respectively. i.e., high altruism if answer > 75). The second quantitative question asks: *How many people know that you commit time to charitable purposes?* Finally, we did PCA analysis, which includes both qualitative questions and quantitative questions by which we constructed the index2. The first component of index2 explains 64.06% of the total variance. We replicated our models for promise to volunteer (columns 5 and 6 in Table C.4) and reneging on a promise (columns 7 and 8 in Table C.5) using index2. The results are presented in Table C.7.

We do not see much difference between the results here and those in previous findings where we used the index constructed by qualitative questions as mediator. We list the empirical results applying the new index as follows. For promise to volunteer: different from previous findings, *Conscientiousness* is now inconsistently mediated (Path $a_1b = 0.062$, 95% CI = [0.005, 0.119], Sobel test p-value = 0.032), and *Agreeableness* now has statistically significant direct effect on making a promise to volunteer. Similar to previous findings, here we find *Agreeableness* is still inconsistently mediated (path $a_1b = 0.256$, 95% CI = [0.124, 0.387], Sobel test p-value < 0.001), and *VolExp* still has significant direct effect on the dependent variable. For reneging on a promise: *Agreeableness* is still inconsistently mediated (Path a_1b = -0.586, 95% CI = [-1.037, -0.136], Sobel test p-value = 0.007), and *Agreeableness*

$\times$ *VolExp* is still inconsistently mediated (Path $a_3b = 0.283$, 95% CI = [0.013, 0.553], Sobel test p-value = 0.043). Finally, AIC in the regressions using index2 as mediator do not show much difference from the previous results.

Volunteering Later vs. Volunteering Now:

As mentioned in Section 3.3.1, we compared the difference between the indirect effect on promise to volunteer with the effect on volunteering now via the altruistic self-concept using additional 134 participants. The results for using volunteering now as the dependent variable is presented in Table C.8. Column (1) reports the total effect, while columns (2) and (3) report indirect and effect effects similar to Table 3.1, and columns (4) and (5) report indirect and effect effects with consideration of moderator and confounder similar to columns (5) and (6) in Table C.4. Total effect and direct effect illustrate *Agreeableness* positively contributes to the likelihood of volunteering now (although not statistically significant after applying the Benjamini-Hochberg procedure, p-value = 0.068).

C.2 Moderated Mediation

In our experiment, ignoring a potential counfounder in the analysis could produce a spurious relationship between personality traits and the willingness to make and keep a promise to volunteer. Unlike a mediator, a confounder can differentially affect and be affected by the independent variable, and the effect of the confounder does not change depending on its value (MacKinnon et al. 2000).

As illustrated by Hayes (2015, 2018); Edwards and Lambert (2007); Preacher et al. (2007), there could exist two stages of moderated mediation. In the first stage of the model, the effects of independent variables (X) on the mediator (M) is moderated by the moderator (W). Similarly, in the second stage of the model, the effects of the mediator (M) on the dependent variable (Y) is moderated by the moderator (W). It is possible that both/either/none of the above stages exist. Figure C.1 and Figure C.2 provide the pathway of the moderated mediation model we have tested.

The statistical forms of the moderated mediation model Stage 1 ($X_i \rightarrow M$) could be expressed as:

$$\begin{cases} M = \beta_0 + a_1 X_i + a_2 W_{1i} + a_3 X_i W_{1i} + \epsilon_M \\ Y = \beta_1 + c' X_i + bM + \epsilon_Y \end{cases}$$

Now, the indirect effect of X on Y consists of the effect from path ab as well as the effect from the moderator. The sum equals to $a_1 b + a_3 b W_1$. This means the mechanism's effect of independent variables on the mediator changes with the size of the moderator.

FIGURE C.1: MEDIATION ANALYSIS W/ MODERATOR: STAGE 1

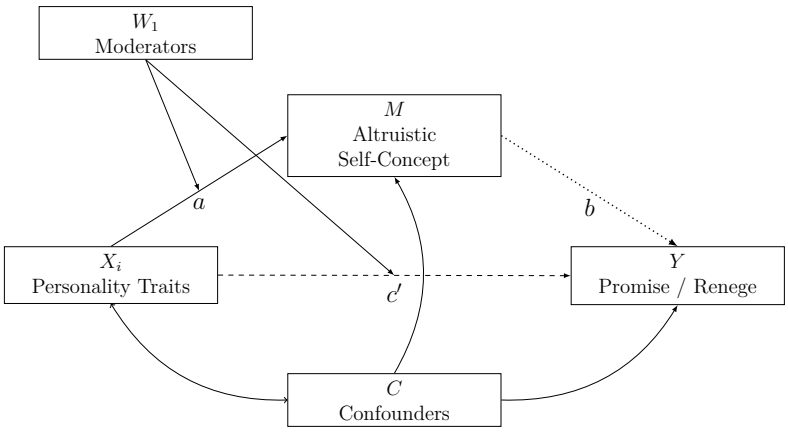
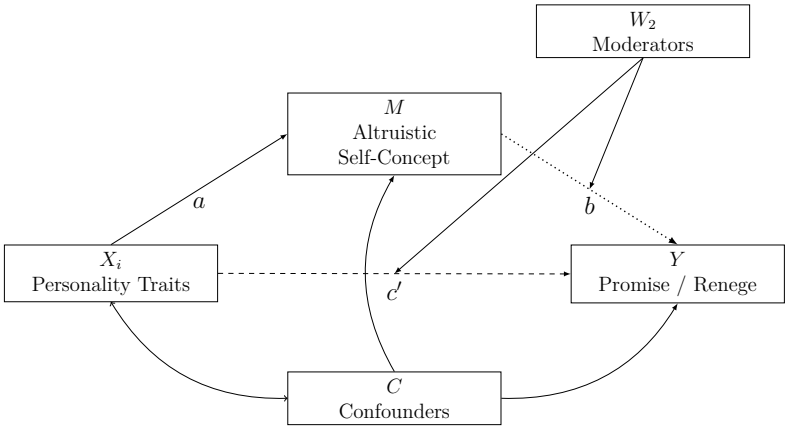


FIGURE C.2: MEDIATION ANALYSIS W/ MODERATOR: STAGE 2



Similarly, the statistical forms of the moderated mediation model Stage 2 ($M \rightarrow Y$) could be expressed as:

$$\begin{cases} M = \beta_0' + aX_i + \epsilon_M' \\ Y = \beta_1' + c'X_i + b_1 M + b_2 W_2 + b_3 M W_2 + \epsilon_Y' \end{cases}$$

Accordingly, the indirect effect of X on Y now turns to be $ab_1 + ab_3 W_2$. This means the mechanism's effect of the mediator on the dependent variable changes with the size of the moderator.

C.3 Tables

Table C.1: Personality Traits and Prosocial Behaviors: Without Mediator(s)

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Panel A: Dictator Game					
Becker et al. (2012)	+	-		+	+
Ben-Ner et al. (2004a)				+	
Ben-Ner et al. (2004b)	+				
Ben-Ner et al. (2008)			-	+	
Ben-Ner and Kramer (2011)		-			+
Brocklebank et al. (2011)	+		-		-
Panel B: Public Goods Game					
Kurzban and Houser (2001)					
Koole (2001)			-	+	
Volk et al. (2011)				+	
Volk et al. (2012)				+	
Panel C: Meta-Analysis					
Kline et al. (2019)	+			+	
Panel D: Non-Game Findings					
Yarkoni et al. (2015)				+	
Brown and Taylor (2015)	+	-		+	-
Pinazo et al. (2016)				+	+
Bekkers (2005, 2010)		-	+		

Note: A summary of highly cited studies on this topic is shown in this table. The table describes whether researchers find a positive (+), negative (-), or no correlation (empty cell) between the Big Five personality traits and prosocial behaviors. Only statistically significant results are reported in this table. Most studies use dictator games and public goods games, so we report the results of a few widely cited literature in Panel A and B, respectively. We report results in Meta-Analysis in Panel C, which includes dictator games, public goods games, prisoner's dilemma, common-pool resource dilemma, ultimatum games, trust games, and Myers-Briggs Type Indicator. Other forms of eliciting prosocial behaviors are listed in Panel D: Yarkoni et al. (2015): real monetary donation. Brown and Taylor (2015): interviews responses about monetary donated and hours volunteered before. Pinazo et al. (2016): six yes-no questions ask about donating, promoting social justice, and volunteering behavior. Bekkers (2005, 2010): questionnaire about how an individual would contribute his/her time or money to a non-profit organization.

Table C.2: Personality Traits and Prosocial Behaviors: With Mediator(s)

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Mediator(s)
Habashi et al. (2016)	+			+	+	Empathic concern, Personal distress
Graziano et al. (2007)				+		Empathic concern
Okun (2007)			+			Social capital
Sun et al. (2019)				+		Social self-efficacy
Caprara et al. (2012)				+		Empathic self-efficacy
Guo et al. (2018)					-	Social self-efficacy, Perspective taking, Empathic concern
Carlo et al. (2005)			+	+		Prosocial value motive
Hill (2016)				+		SRAS

Note: Forms of prosocial behaviors in this Table: Habashi et al. (2016) and Graziano et al. (2007): participants are required to select the number of hours they are willing to volunteer. Okun (2007): participants report their volunteering experience during the past month. Sun et al. (2019): a Chinese version of the Self-Report Altruism Scale Distinguished by the Participants' Prosocial Behaviors (SRAS-DR; Oda et al., 2013). Questions measure the prosociality toward family member, friends/acquaintances, and strangers). Caprara et al. (2012): participants rate their prosociality on a 16-item scale. Guo et al. (2018): adopt questions from Carlo et al. (2003), which measure six types of prosocial behaviors in public, anonymous, dire, emotional, compliant, and altruism. Carlo et al. (2005): four items about volunteering experience and willingness to volunteer in the future. Hill (2016): attitudes, intentions to register, and donor status in organ donation.

Table C.3: Timeline

Day 1

- Demographics (e.g., age, gender, ethnicity, and religion)
- Behavioral tendencies (e.g., exercise, smoking, drinking)
- Mensa Quiz
- Time Preference Quantitative Measure
- Locus of Control
- **Decision of making a promise to volunteer**

Day 4

- Big Five
- Altruistic Self-Concept:
  - (1) Qualitative Measure:
    - willingness to share compared to others
    - willingness to share towards different recipients
    - self-evaluation of altruistic tendency
  - (2) Quantitative Measure:
    - dictator game
    - observability of volunteering
- Time Preference Qualitative Measure
- For those who promised to volunteer on Day 1: **Decision of keeping or reneging on a promise to volunteer**
- Volunteer survey (if promise to volunteer is kept)

Table C.4: Moderated Mediation - Promise to Volunteer - Stage 1

	(1) Index	(2) Promise	(3) Index	(4) Promise	(5) Index	(6) Promise	(7) Index	(8) Promise	(9) Index	(10) Promise
Index	NA	**0.468*****	NA	**0.452****	NA	**0.459****	NA	**0.476*****	NA	**0.442****
Conscientiousness	0.138	−0.342	0.063	−0.347	0.167	−0.293	0.101	−0.052	0.019	−0.189
Neuroticism	0.072	−0.155	0.152	0.194	0.138	−0.014	0.151	0.087	0.142	0.204
Openness	0.225*	0.159	0.343*	0.353	0.189*	0.073	0.220	0.260	0.313	0.438
Extraversion	−0.116	0.102	−0.189	0.075	−0.103	0.123	−0.159	0.089	−0.210	0.046
Agreeableness	**0.960*****	−0.410*	**0.989*****	−0.274	**0.964*****	0.404*	**1.046*****	0.325	**1.038*****	−0.228
Female	0.319	0.732*	0.684	1.959					0.612	1.235
VolExp					**0.356***	**1.113*****	0.416	3.325**	0.125	2.983*
Conscientiousness × Female			0.222	0.064					0.201	0.300
Neuroticism × Female			−0.098	−0.415					−0.091	−0.343
Openness × Female			−0.392	−0.498					−0.392	−0.461
Extraversion × Female			0.119	0.060					0.085	0.037
Agreeableness × Female			−0.043	−0.218					0.014	−0.178
Conscientiousness × VolExp							0.232	−0.598	0.190	−0.642
Neuroticism × VolExp							−0.012	−0.102	0.019	−0.039
Openness × VolExp							−0.105	−0.535	0.076	−0.356
Extraversion × VolExp							0.117	0.015	0.090	0.003
Agreeableness × VolExp							−0.240	−0.152	−0.264	−0.177
N. Obs	251	251	251	251	251	251	251	251	251	251
(Pseudo) R-Square	0.446	0.0.79	0.459	0.090	0.449	0.108	0.457	0.121	0.473	0.140
AIC	1120.839		1131.021		1109.644		1121.692		1132.223	

$*$ p $<$ 0.05; $**$ p $<$ 0.01; $***$ p $<$ 0.001. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Table C.5: Moderated Mediation - Reneging on a Promise - Stage 1

	(1) Index	(2) Renege	(3) Index	(4) Renege	(5) Index	(6) Renege	(7) Index	(8) Renege	(9) Index	(10) Renege
Index	NA	−0.557*	NA	−0.668*	NA	−0.702*	NA	**-0.898****	NA	−1.025*
Conscientiousness	0.183	−0.587	−0.040	−1.054	0.250	−0.572	−0.089	−1.150	−0.136	−1.977**
Neuroticism	0.214	−0.251	0.227	−0.194	0.303	−0.505	0.555*	−0.338	0.388	0.227
Openness	**0.383****	0.598	0.393	1.156	0.315*	0.789*	0.043	0.710	0.144	1.268
Extraversion	−0.199	0.284	−0.584*	0.312	−0.155	0.123	0.051	−0.071	−0.459	0.147
Agreeableness	**1.025*****	0.206	**0.984*****	0.315	1.070***	0.172	**1.567*****	0.628	**1.345*****	0.879
Female	0.528*	−1.110	−0.385	0.273					−1.671	0.873
VolExp					0.371	0.384	−0.179	−0.988	1.307	−1.614
Conscientiousness × Female			0.482	1.012					0.282	0.571
Neuroticism × Female			−0.057	−0.034					0.351	−0.072
Openness × Female			−0.053	−1.334					0.058	−1.831
Extraversion × Female			0.435*	−0.015					**0.702****	0.224
Agreeableness × Female			−0.025	0.038					0.602	0.312
Conscientiousness × VolExp							0.586*	0.982	0.453	1.848
Neuroticism × VolExp							−0.273	−0.242	−0.495	−0.444
Openness × VolExp							0.520	0.302	0.221	0.320
Extraversion × VolExp							−0.049	0.415	−0.257	0.226
Agreeableness × VolExp							**-0.873****	−0.742	**-1.239*****	−0.906
N. Obs	92	92	92	92	92	92	92	92	92	92
(Pseudo) R-Square	0.525	0.210	0.578	0.228	0.513	0.183	0.572	0.218	0.650	0.285
AIC	362.019		369.536		366.299		371.730		371.981	

$*$ $p < 0.05$; $**$ $p < 0.01$; $***$ $p < 0.001$. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Table C.6: Moderated Mediation - Promise to Volunteer and Reneging on a Promise - Stage 2

	(1) Index	(2) Promise	(3) Index	(4) Promise	(5) Index	(6) Renege	(7) Index	(8) Renege
Index	NA	0.333*	NA	0.442*	NA	−0.694*	NA	−0.819*
Conscientiousness	0.138	−0.341*	0.167	−0.295	0.183	−0.626	0.250	−0.611
Neuroticism	0.072	−0.146	0.138	−0.015	0.214	−0.234	0.303	−0.490
Openness	0.225*	0.157	0.189*	0.071	**0.383****	0.612	0.315*	0.789*
Extraversion	−0.116	0.090	−0.103	0.121	−0.199	0.188	−0.155	0.091
Agreeableness	**0.960*****	−0.420*	**0.964*****	−0.402*	**1.025*****	0.205	1.070***	0.224
Female	0.319	0.680*			0.528*	−1.082		
VolExp			**0.356***	**1.106*****			0.371	0.431
Index × Female		0.289				0.360		
Index × VolExp				0.037				0.182
N. Obs	251	251	251	251	92	92	92	92
(Pseudo) R-Square	0.446	0.085	0.449	0.108	0.525	0.217	0.217	0.223

Note: $*$ $p < 0.05$; $**$ $p < 0.01$; $***$ $p < 0.001$. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Table C.7: Mediation Analysis - Index2

	(1) Index	(2) Promise	(3) Index	(4) Renege
Index	NA	**0.539*****	NA	**-0.922****
Conscientiousness	**0.228****	−0.347	0.021	−1.059
Neuroticism	0.122	−0.016	**0.574***	−0.348
Openness	**0.208***	0.054	0.094	0.824
Extraversion	−0.065	0.111	0.047	−0.137
Agreeableness	**0.920*****	**-0.457***	**1.438*****	0.492
VolExp	**0.745*****	**0.907****	0.454	−0.494
Conscientiousness × VolExp			0.558*	0.944
Neuroticism × VolExp			−0.356	−0.354
Openness × VolExp			0.343	0.200
Extraversion × VolExp			−0.021	0.473
Agreeableness × VolExp			**-0.788****	−0.744
N. Obs	251	251	92	92
(Pseudo) R-Square	0.488	0.122	0.593	0.219
AIC	1107.880		371.250	

$*$ p $<$ 0.05; $**$ p $<$ 0.01; $***$ p $<$ 0.001. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Table C.8: Mediation Analysis - Volunteering Now

	Path c (1) Volunteer Now	Path a (2) Index	Path b & c' (3) Volunteer Now	Path a (4) Index	Path b & c' (5) Volunteer Now
Index	NA	NA	**0.424***	NA	0.375*
Conscientiousness	−0.067	0.167	−0.102	0.192	−0.007
Neuroticism	0.418	0.153	0.379	0.143	0.378
Openness	**0.488****	0.171	0.418	0.130	0.307
Extraversion	−0.023	0.251*	−0.139	0.222	−0.213
Agreeableness	**0.913*****	**0.775*****	**0.668***	**0.743*****	0.678*
VolExp				0.363	1.145**
N. Obs	134	134	134	134	134
(Pseudo) R-Square	0.111	0.374	0.144	0.360	0.185

$*$ p $<$ 0.05; $**$ p $<$ 0.01; $***$ p $<$ 0.00. The bold numbers highlight the variables whose p-values are strictly lower than 0.05 after adjusting for multiple hypotheses testing using the Benjamini-Hochberg procedure.

Bibliography

[1] Abeler, Johannes, et al. "Reference points and effort provision." *American Economic Review* 101.2 (2011): 470-92.

[2] Abraham, Katharine G., et al. "Measuring the Gig Economy: Current Knowledge and Open Issues." Working Paper (2018).

[3] Accounting Principals. Holiday bonus and hiring survey. 2017.

[4] Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. "Personality psychology and economics." In *Handbook of the Economics of Education*, vol. 4, pp. 1-181. Elsevier, 2011.

[5] Andreoni, James. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of Political Economy* 97, no. 6 (1989): 1447-1458.

[6] Andreoni, James. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The Economic Journal* 100, no. 401 (1990): 464-477.

[7] Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *The Quarterly Journal of Economics*, 116(1), 293-312.

[8] Andreoni, James, and B. Douglas Bernheim. "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects." *Econometrica* 77, no. 5 (2009): 1607-1636.

[9] Andreoni, James, and Marta Serra-Garcia. "Time-inconsistent charitable giving." Working Paper (2016).

[10] Andreoni, James, Justin M. Rao, and Hannah Trachtman. "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving." *Journal of Political Economy* 125, no. 3 (2017): 625-653.

[11] Anik, Lalin, et al. "pro-social bonuses increase employee satisfaction and team performance." *PloS one* 8.9 (2013): e75509.

[12] Arechar, Antonio A., Simon Gächter, and Lucas Molleman. “Conducting interactive experiments online." *Experimental Economics* 21, no. 1 (2018): 99-131.

[13] Ariely, D., & Norton, M. I. (2008). How actions create–not just reveal–preferences. *Trends in cognitive sciences*, 12(1), 13-16.

[14] Augenblick, Ned, Muriel Niederle, and Charles Sprenger. “Working over time: Dynamic inconsistency in real effort tasks." *The Quarterly Journal of Economics* 130, no. 3 (2015): 1067-1115.

[15] Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of personality and social psychology*, 97(3), 533.

[16] Banuri, Sheheryar, and Philip Keefer. "Pro-social motivation, effort and the call to public service." *European Economic Review* 83 (2016): 139-164.

[17] Barrett, T. S., Lockhart, G., & Cruz, R. A. (2019). Marginal Mediation Analysis: A Practical Statistical Framework for Interpretable Mediation Effects. Working Paper.

[18] Barrick, M. R., & Mount, M. K. (2000). Select on conscientiousness and emotional stability. *Handbook of principles of organizational behavior*, 15, 28.

[19] Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics*, 4(1), 453-478.

[20] Bekkers, R. (2005). Participation in voluntary associations: Relations with resources, personality, and political values. *Political psychology*, 26(3), 439-454.

[21] Bekkers, R. (2006). Traditional and health-related philanthropy: The role of resources and personality. *Social psychology quarterly*, 69(4), 349-366.

[22] Bekkers, R. (2010). Who gives what and when? A scenario study of intentions to give time and money. *Social Science Research*, 39(3), 369-381.

[23] Bekkers, R., & Wiepking, P. (2011). A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving. *Nonprofit and voluntary sector quarterly*, 40(5), 924-973.

[24] Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American economic review*, 96(5), 1652-1678.

[25] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.

[26] Benndorf, Volker, Holger A. Rau, and Christian Sölch. "Gender differences in motivational crowding out of work performance." *Economic Inquiry* 57.1 (2019): 206-226.

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

[28] Ben-Ner, A., Putterman, L., Kong, F., & Magan, D. (2004). Reciprocity in a two-part dictator game. *Journal of Economic Behavior & Organization*, 53(3), 333-352.

[29] Ben-Ner, A., Kramer, A., & Levy, O. (2008). Economic and hypothetical dictator game experiments: Incentive effects at the individual level. *The Journal of Socio-Economics*, 37(5), 1775-1784.

[30] Ben-Ner, A., & Kramer, A. (2011). Personality and altruism in the dictator game: Relationship to giving to kin, collaborators, competitors, and neutrals. *Personality and Individual Differences*, 51(3), 216-221.

[31] Berger, Leslie and Guo, Lan and Presslee, Adam. "Better to Give than to Receive: The Positive Effect of pro-social Rewards on Goal Commitment and Performance." Working Paper (2018).

[32] Biderman, M. D., Nguyen, N. T., & Sebren, J. (2008). Time-on-task mediates the conscientiousness–performance relationship. *Personality and Individual Differences*, 44(4), 887-897.

[33] Bloom, P. (2017). Against empathy: The case for rational compassion. Random House.

[34] Borghans, Lex, James J. Heckman, Bart HH Golsteyn, and Huub Meijers. "Gender differences in risk aversion and ambiguity aversion." *Journal of the European Economic Association* 7, no. 2-3 (2009): 649-658.

[35] Breman, Anna. “Give more tomorrow: Two field experiments on altruism and intertemporal choice." *Journal of Public Economics* 95, no. 11-12 (2011): 1349-1357.

[36] Brocklebank, S., Lewis, G. J., & Bates, T. C. (2011). Personality accounts for stable preferences and expectations across a range of simple games. *Personality and Individual Differences*, 51(8), 881-886.

[37] Brown, S., & Taylor, K. (2015). Charitable behaviour and the big five personality traits: evidence from UK panel data. Working Paper.

[38] Brown, Alexander L., Jonathan Meer, and J. Forrest Williams. “Why do people volunteer? An experimental analysis of preferences for time donations." *Management Science* 65, no. 4 (2018): 1455-1468.

[39] Bryan, Gharad, Dean Karlan, and Scott Nelson. "Commitment devices." *Annual Review of Economics* 2.1 (2010): 671-698.

[40] Burbano, Vanessa C. "Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces." *Organization Science* 27.4 (2016): 1010-1028.

[41] Burnett, P. C. (1991). Decision-making style and self-concept. *Australian Psychologist*, 26(1), 55-58.

[42] Capra, C. Monica, Bing Jiang, Jan B. Engelmann, and Gregory S. Berns. “Can personality type explain heterogeneity in probability distortions?." *Journal of Neuroscience, Psychology, and Economics* 6, no. 3 (2013): 151.

[43] Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of personality and social psychology*, 102(6), 1289.

[44] Carlo, G., Hausmann, A., Christiansen, S., & Randall, B. A. (2003). Sociocognitive and behavioral correlates of a measure of prosocial tendencies for adolescents. *The journal of early adolescence*, 23(1), 107-134.

[45] Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion and prosocial value motivation. *Personality and individual differences*, 38(6), 1293-1305.

[46] Carpenter, Jeffrey, and Erick Gong. "Motivating Agents: How much does the mission matter?." *Journal of Labor Economics* 34.1 (2016): 211-236.

[47] Cassar, Lea, and Stephan Meier. Intentions for doing good matter for doing well: The (negative) signaling value of prosocial incentives. No. w24109. National Bureau of Economic Research, 2018.

[48] Charness, Gary, Ramon Cobo-Reyes, and Angela Sanchez. "The effect of charitable giving on workers' performance: Experimental evidence." *Journal of Economic Behavior & Organization* 131 (2016): 61-74.

[49] Charness, Gary, Uri Gneezy, and Austin Henderson. "Experimental methods: Measuring effort in economics experiments." *Journal of Economic Behavior & Organization* 149 (2018): 74-87.

[50] Chou, Eileen Y., Dennis Y. Hsu, and Eileen Hernon. "From slacktivism to activism: Improving the commitment power of e-pledges for prosocial causes." *PloS one* 15.4 (2020): e0231314.

[51] Chuan, Amanda, and Anya Savikhin Samek. Chuan, Amanda, and Anya Savikhin Samek. ""Feel the Warmth" glow: A field experiment on manipulating the act of giving." *Journal of Economic Behavior & Organization* 108 (2014): 198-211.

[52] Clary, E. G., & Snyder, M. (1999). The motivations to volunteer: Theoretical and practical considerations. *Current directions in psychological science*, 8(5), 156-159.

[53] Crumpler, Heidi, and Philip J. Grossman. "An experimental test of warm glow giving." *Journal of Public Economics* 92, no. 5-6 (2008): 1011-1021.

[54] Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11-15.

[55] Corgnet, Brice, Joaquín Gómez-Miñambres, and Roberto Hernán-Gonzalez. "Goal setting and monetary incentives: When large stakes are not enough." *Management Science* 61.12 (2015): 2926-2944.

[56] Costa, Paul T., and Robert R. McCrae. "Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)." Odessa, FL: *Psychological Assessment Resources* 61 (1992).

[57] Cotterill, Sarah, Peter John, and Liz Richardson. "The impact of a pledge request and the promise of publicity: a randomized controlled trial of charitable donations." *Social Science Quarterly* 94.1 (2013): 200-216.

[58] Cox, J. C., & Deck, C. A. (2006). When are women more generous than men?. *Economic Inquiry*, 44(4), 587-598.

[59] Cunningham, John A., Alexandra Godinho, and Vladyslav Kushnir. "Can Amazon's Mechanical Turk be used to recruit participants for internet intervention trials? A pilot study involving a randomized controlled trial of a brief online intervention for hazardous alcohol use." *Internet interventions* 10 (2017): 12-16.

[60] Dalton, Patricio S., V. Gonzalez, and Charles N. Noussair. Paying with self-chosen goals: Incentives and gender differences. No. 2015-021. CentER Discussion Paper Series, 2015.

[61] Damgåard, Mette Trier, and Christina Gravert. "Now or never! The effect of deadlines on charitable giving: Evidence from two natural field experiments." *Journal of Behavioral and Experimental Economics* 66 (2017): 78-87.

[62] DellaVigna, Stefano, John A. List, and Ulrike Malmendier. "Testing for altruism and social pressure in charitable giving." *The Quarterly Journal of Economics* 127, no. 1 (2012): 1-56.

[63] DellaVigna, Stefano, et al. Estimating social preferences and gift exchange at work. No. w22043. National Bureau of Economic Research, 2016.

[64] DellaVigna, Stefano, and Devin Pope. "What motivates effort? Evidence and expert forecasts." *The Review of Economic Studies* 85.2 (2017): 1029-1069.

[65] Dreber, Anna, Drew Fudenberg, David K. Levine, and David G. Rand. “Self-control, social preferences and the effect of delayed payments." Working Paper (2016).

[66] Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological bulletin*, 100(3), 283.

[67] Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological methods*, 12(1), 1.

[68] Erkal, Nisvan, Lata Gangadharan, and Nikos Nikiforakis. "Relative earnings and giving in a real-effort experiment." *American Economic Review* 101.7 (2011): 3330-48.

[69] Evren, Özgür, and Stefania Minardi. “Warm-glow Giving and Freedom to be Selfish." *The Economic Journal* 127, no. 603 (2017): 1381-1409.

[70] Falk, Armin, Anke Becker, Thomas J. Dohmen, David Huffman, and Uwe Sunde. “The preference survey module: A validated instrument for measuring risk, time, and social preferences." Working Paper (2016).

[71] Farrell, Diana, and Fiona Greig. "The online platform economy: Has growth peaked?." Working Paper (2017).

[72] Fosgaard, Toke, and Adriaan R. Soetevent. “Promises undone: How committed pledges impact donations to charity." Working Paper (2018).

[73] Furnham, A., Treglown, L., Hyde, G., & Trickey, G. (2016). The bright and dark side of altruism: Demographic, personality traits, and disorders associated with altruism. *Journal of Business Ethics*, 134(3), 359-368.

[74] Goerg, Sebastian J., and Sebastian Kube. "Goals (th) at work–goals, monetary incentives, and workers’ performance." MPI Collective Goods Preprint 2012/19 (2012).

[75] Gómez-Miñambres, Joaquín. "Motivation through goal setting." *Journal of Economic Psychology* 33.6 (2012): 1223-1239.

[76] Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person × situation perspective. *Journal of personality and social psychology*, 93(4), 583.

[77] Grosswirth, Marvin, and Abbie F. Salny. The Mensa genius quiz. Da Capo Press, 1983.

[78] Guo, Q., Sun, P., & Li, L. (2018). Why neurotic individuals are less prosocial? A multiple mediation analysis regarding related mechanisms. *Personality and Individual Differences*, 128, 55-61.

[79] Habashi, M. M., Graziano, W. G., & Hoover, A. E. (2016). Searching for the prosocial personality: A Big Five approach to linking personality and prosocial behavior. *Personality and Social Psychology Bulletin*, 42(9), 1177-1192.

[80] Hara, Kotaro, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham. "A data-driven analysis of workers' earnings on amazon mechanical turk." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 449. ACM, 2018.

[81] Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate behavioral research*, 50(1), 1-22.

[82] Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, 85(1), 4-40.

[83] Heckman, J. J., Jagelka, T., & Kautz, T. D. (2019). Some Contributions of Economics to the Study of Personality (No. w26459). National Bureau of Economic Research. Working Paper.

[84] Hill, E. M. (2016). Posthumous organ donation attitudes, intentions to donate, and organ donor status: Examining the role of the big five personality dimensions and altruism. *Personality and Individual Differences*, 88, 182-186.

[85] Hirsh, J. B. (2010). Personality and environmental concern. *Journal of Environmental Psychology*, 30(2), 245-248.

[86] Hossain, Tanjim, and King King Li. "Crowding out in the labor market: A prosocial setting is necessary." *Management Science* 60.5 (2014): 1148-1160.

[87] Hotchkiss, R. B., Fottler, M. D., & Unruh, L. (2008, August). Valuing volunteers: the impact of volunteerism on hospital performance. *Academy of Management Proceedings* (Vol. 2008, No. 1, pp. 1-6). Briarcliff Manor, NY 10510: Academy of Management.

[88] Hotchkiss, R. B., Unruh, L., & Fottler, M. D. (2014). The role, measurement, and impact of volunteerism in hospitals. *Nonprofit and Voluntary Sector Quarterly*, 43(6), 1111-1128.

[89] Imas, Alex. "Working for the "warm glow": On the benefits and limits of prosocial incentives." *Journal of Public Economics* 114 (2014): 14-18.

[90] Imbens, G. W. (2020). Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature*, 58(4), 1129-1179.

[91] Jenkinson, Caroline E., Andy P. Dickens, Kerry Jones, Jo Thompson-Coon, Rod S. Taylor, Morwenna Rogers, Clare L. Bambra, Iain Lang, and Suzanne H. Richards. "Is volunteering a public health intervention? A systematic review and meta-analysis of the health and survival of volunteers." *BMC public health* 13, no. 1 (2013): 773.

[92] Jiang, Bing, and C. Monica Capra. "Are (active) entrepreneurs a different breed?." *Managerial and Decision Economics* 39, no. 6 (2018): 613-628.

[93] John, Oliver P., and Sanjay Srivastava. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." *Handbook of Personality: Theory and Research* 2, no. 1999 (1999): 102-138.

[94] Klein, Howard J., and Sunhee Lee. "The effects of personality on learning: The mediating role of goal setting." *Human Performance* 19.1 (2006): 43-66.

[95] Khan, Uzma, Kelly Goldsmith, and Ravi Dhar. "When Does Altruism Trump Self-Interest? The Moderating Role of Affect in Extrinsic Incentives." *Journal of the Association for Consumer Research* 5.1 (2020): 000-000.

[96] King, H. R., Jackson, J. J., Morrow-Howell, N., & Oltmanns, T. F. (2014). Personality accounts for the connection between volunteering and health. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 70(5), 691-697.

[97] Kitzmueller, Markus, and Jay Shimshack. "Economic perspectives on corporate social responsibility." *Journal of Economic Literature* 50.1 (2012): 51-84.

[98] Kentner, N., Lange, C., Reifschneider, E., & Takacs, A. (2003). The cost and benefits of volunteers. East Lansing, MI: Michigan State University Extension.

[99] Koole, S. L., Jager, W., van den Berg, A. E., Vlek, C. A., & Hofstee, W. K. (2001). On the social nature of personality: Effects of extraversion, agreeableness, and feedback about collective resource use on cooperation in a resource dilemma. *Personality and Social Psychology Bulletin*, 27(3), 289-301.

[100] Kline, R., Bankert, A., Levitan, L., & Kraft, P. (2017). Personality and Prosocial Behavior: A Multilevel Meta-Analysis. *Political Science Research and Methods*, 1-18.

[101] Knowles, Stephen, and Maroš Servátka. "Transaction costs, the opportunity cost of time and procrastination in charitable giving." *Journal of Public Economics* 125 (2015): 54-63.

[102] Knowles, Stephen, Maroš Servátka, and Trudy Sullivan. "Deadlines, procrastination, and inattention in charitable tasks: A field experiment." Working Paper (2018).

[103] Kőszegi, Botond, and Matthew Rabin. "A model of reference-dependent preferences." *The Quarterly Journal of Economics* 121.4 (2006): 1133-1165.

[104] Kőszegi, Botond, and Matthew Rabin. "Reference-dependent risk attitudes." *American Economic Review* 97.4 (2007): 1047-1073.

[105] Kurzban, R., & Houser, D. (2001). Individual differences in cooperation in a circular public goods game. *European Journal of Personality*, 15(S1), S37-S52.

[106] Lacetera, Nicola, Mario Macis, and Angelo Mele. "Viral Altruism? Charitable Giving and Social Contagion in Online Networks." *Sociological Science* 3 (2016).

[107] Lee, H., Herbert, R. D., & McAuley, J. H. (2019). Mediation analysis. *Jama*, 321(7), 697-698.

[108] Li, King King. "The Hidden Cost of Using CSR to Motivate Workers." Working Paper (2019).

[109] Lilley, Andrew, and Robert Slonim. "The price of warm glow." *Journal of Public Economics* 114 (2014): 58-74.

[110] List, John A., and Fatemeh Momeni. When corporate social responsibility backfires: Theory and evidence from a natural field experiment. No. w24169. National Bureau of Economic Research, 2017.

[111] Locke, Edwin A., and Gary P. Latham. A theory of goal setting & task performance. Prentice-Hall, Inc, 1990.

[112] Lorko, Matej, Maroš Servátka, and Le Zhang. "Anchoring in project duration estimation." *Journal of Economic Behavior & Organization* 162 (2019): 49-65.

[113] MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention science*, 1(4), 173-181.

[114] MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, 58, 593-614.

[115] Meer, Jonathan. "Brother, can you spare a dime? Peer pressure in charitable solicitation." *Journal of Public Economics* 95, no. 7-8 (2011): 926-941.

[116] Meer, Jonathan, and Oren Rigbi. "The effects of transactions costs and social distance: evidence from a field experiment." *The BE Journal of Economic Analysis & Policy* 13, no. 1 (2013): 271-296.

[117] Meyer, R. D., Dalal, R. S., & Bonaccio, S. (2009). A meta-analytic investigation into the moderating effects of situational strength on the conscientiousness–performance relationship. *Journal of Organizational Behavior*, 30(8), 1077-1102.

[118] Meyer, Christian, and Egon Tripodi. "Image Concerns in Pledges to Give Blood: Evidence from a Field Experiment." Working Paper (2018)

[119] Meyvis, Tom, Aronte Bennett, and Daniel M. Oppenheimer. "Precommitment to charity." *The science of giving: Experimental approaches to the study of charity* (2011): 35-48.

[120] Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., & Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, 75(1), 73-79.

[121] Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in personnel and human resources management*, 13(3), 153-200.

[122] Müller, J., & Schwieren, C. (2019). Big Five personality factors in the Trust Game. *Journal of Business Economics*, 1-19.

[123] O'Brien, Ed, and Samantha Kassirer. "People are slow to adapt to the warm glow of giving." *Psychological science* 30.2 (2019): 193-204.

[124] Oda, R., Shibata, A., Kiyonari, T., Takeda, M., & Matsumoto-Oda, A. (2013). Sexually dimorphic preference for altruism in the opposite sex according to recipient. *British Journal of Psychology*, 104(4), 577-584.

[125] Oda, R., Machii, W., Takagi, S., Kato, Y., Takeda, M., Kiyonari, T., Fukukawa, Y., & Hiraishi, K. (2014). Personality and altruism in daily life. *Personality and Individual Differences*, 56, 206-209.

[126] Okun, M. A., Pugliese, J., & Rook, K. S. (2007). Unpacking the relation between extraversion and volunteering in later life: The role of social capital. *Personality and individual differences*, 42(8), 1467-1477.

[127] Pavalache-Ilie, M., & Cazan, A. M. (2018). Personality correlates of pro-environmental attitudes. *International journal of environmental health research*, 28(1), 71-78.

[128] Pearl, J., & Mackenzie, D. (2015). The book of why: the new science of cause and effect. Basic Books.

[129] Peer, Eyal, Joachim Vosgerau, and Alessandro Acquisti. “Reputation as a sufficient condition for data quality on Amazon Mechanical Turk." *Behavior Research Methods* 46, no. 4 (2014): 1023-1031.

[130] Peifer, Jared L. "The economics and sociology of religious giving: Instrumental rationality or communal bonding?." *Social Forces* 88.4 (2010): 1569-1594.

[131] Perugini, M., & Leone, L. (2009). Implicit self-concept and moral action. *Journal of Research in Personality*, 43(5), 747-754.

[132] Piliavin, J. A., & Callero, P. L. (1991). Giving blood: the development of an altruistic identity. Johns Hopkins University Press.

[133] Pinazo, D., Barrós-Loscertales, A., Peris, R., Ventura-Campos, N., & Ávila, C. (2016). Neuroticism predisposes to donation more than agreeableness: An fMRI study. *Journal of Neuroscience, Psychology, and Economics*, 9(2), 100.

[134] Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, 42(1), 185-227.

[135] Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891.

[136] Presslee, Adam, Thomas W. Vance, and R. Alan Webb. "The effects of reward type on employee goal setting, goal commitment, and performance." *The Accounting Review* 88.5 (2013): 1805-1831.

[137] Qiao, Dandan, et al. "Overcoming the Crowding-Out Effect of Monetary Incentive on Pro-Social Behavior." Working Paper (2017).

[138] Huck, Steffen, and Imran Rasul. “Transactions costs in charitable giving: evidence from two field experiments." *The BE Journal of Economic Analysis & Policy* 10, no. 1 (2010).

[139] Resh, William G., John D. Marvel, and Bo Wen. "The persistence of prosocial work effort as a function of mission match." *Public Administration Review* 78.1 (2018): 116-125.

[140] Rotter, Julian B. "Generalized expectancies for internal versus external control of reinforcement." *Psychological Monographs: General and Applied* 80, no. 1 (1966): 1.

[141] Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and individual differences*, 2(4), 293-302.

[142] Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, 64, 122-137.

[143] Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, 8(4), 379-386.

[144] Schroeder, Juliana, and Ayelet Fishbach. "How to motivate yourself and others? Intended and unintended consequences." *Research in Organizational Behavior* 35 (2015): 123-141.

[145] Schwartz, Daniel, et al. "Opting-in to prosocial incentives." *Organizational Behavior and Human Decision Processes* (2019).

[146] Smith, Sarah, Frank Windmeijer, and Edmund Wright. "Peer effects in charitable giving: Evidence from the (running) field." *The Economic Journal* 125, no. 585 (2015): 1053-1071.

[147] Smith, Jason. "The motivational effects of mission matching: A lab-experimental test of a moderated mediation model." *Public Administration Review* 76.4 (2016): 626-637.

[148] Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of experimental social psychology*, 37(3), 228-243.

[149] Sun, P., Guo, Q., Liu, Z., & Fan, J. (2019). Shyness weakens the agreeableness-prosociality association via social self-efficacy: A moderated-mediation study of Chinese undergraduates. *Frontiers in psychology*, 10, 1084.

[150] Sutan, Angela, et al. "“Facta non verba”: An experiment on pledging and giving." *Journal of Economic Psychology* 65 (2018): 1-15.

[151] Thaler, Richard H. "Mental accounting matters." *Journal of Behavioral decision making* 12.3 (1999): 183-206.

[152] Tonin, Mirco, and Michael Vlassopoulos. “Disentangling the sources of pro-socially motivated effort: A field experiment." *Journal of Public Economics* 94, no. 11-12 (2010): 1086-1092.

[153] Tonin, Mirco, and Michael Vlassopoulos. "Experimental evidence of self-image concerns as motivation for giving." *Journal of Economic Behavior & Organization* 90 (2013): 19-27.

[154] Tonin, Mirco, and Michael Vlassopoulos. "Corporate philanthropy and productivity: Evidence from an online real effort experiment." *Management Science* 61.8 (2015): 1795-1811.

[155] Tversky, Amos, and Daniel Kahneman. "Prospect theory: An analysis of decision under risk." *Econometrica* 47.2 (1979): 263-291.

[156] VandenBos, G. R. (2007). APA dictionary of psychology. American Psychological Association.

[157] Volk, S., Thöni, C., & Ruigrok, W. (2011). Personality, personal values and cooperation preferences in public goods games: A longitudinal study. *Personality and individual Differences*, 50(6), 810-815.

[158] Volk, S., Thöni, C., & Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation preferences. *Journal of Economic Behavior & Organization*, 81(2), 664-676.

[159] Ward, S. J., & King, L. A. (2018). Religion and moral self-image: The contributions of prosocial behavior, socially desirable responding, and personality. *Personality and Individual Differences*, 131, 222-231.

[160] Wells, B. L., DePue, J. D., Buehler, C. J., Lasater, T. M., & Carleton, R. A. (1990). Characteristics of volunteers who deliver health education and promotion: A comparison with organization members and program participants. *Health Education Quarterly*, 17(1), 23-35.

[161] Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of applied psychology*, 87(1), 164.

[162] Wu, George, Chip Heath, and Richard Larrick. "A prospect theory model of goal behavior." Unpublished manuscript (2008).

[163] Yarkoni, T., Ashar, Y. K., & Wager, T. D. (2015). Interactions between donor Agreeableness and recipient characteristics in predicting charitable donation and positive social evaluation. *PeerJ*, 3, e1089.

[164] Young, L., Chakroff, A., & Tom, J. (2012). Doing good leads to more good: The reinforcing power of a moral self-concept. *Review of Philosophy and Psychology*, 3(3), 325-334.